

여성 영업직종간 전기자료 타당도의 안정성*

이 종 구[†] 김 명 소 윤 창 영
대구대학교 심리학과 호서대학교 경영학부 대구대학교 응용심리연구소

본 연구는 하나의 여성 영업직을 대상으로 개발된 전기자료의 가중치를 다른 여성 영업직에 적용했을 때 타당도가 안정적으로 나타나는지를 확인하기 위해 시도되었다. 경험적 점수화 방법을 사용하여 대표본을 가중치 개발집단으로 이용하고 소표본을 교차타당화 집단으로 이용했을 때나 그 역으로 자료를 이용했을 때 모두 가중치 개발집단의 타당도는 물론 교차타당화 집단의 타당도가 유의하게 나타나 한 조직에서 개발된 전기자료의 가중치가 일반화 가능성을 보여주었다. 그러나 가중치 개발집단으로 대표본을 이용했을 때가 소표본을 이용했을 때보다 교차타당화 집단의 타당도 감소가 적응과 동시에 타당도도 더 높게 나타났다. 이 결과는 대규모 조직에서 개발된 전기자료 검사는 소규모 조직의 유사 직종에 적용할 수 있음을 시사해 준다. 또 검사개발과정에서 대표본을 확보하기 위해 여러 조직에 걸쳐 전기자료 검사의 개발이 필요하다는 것을 시사해 준다.

주제어: 전기자료, 타당도 일반화, 경험적 점수화, 여성 영업직

본 연구는 하나의 여성 영업직을 대상으로 개발된 전기자료(biographical data: 일명 biodata) 문항과 채점기준을 가지고 다른 여성 영업직에도 적용할 수 있는지의 여부를 판단하기 위해 시도되었다. 전기자료의 주된 가중치 부여방식인 Strong의 WAB(weighted application blank) 기법에서는 대규모의 자료가 요구된다. 그러나 검사개발 대상 조직의 규모가 작을 경우 동일 직종내

* 익명의 세 분 심사위원의 세심한 조언에 진심으로 감사드립니다.
* 이 논문은 2002년도 대구대학교 학술연구비 지원에 의한 논문임
† 교신저자: 이종구, (712-714) 경북 경산시 진량읍 내리리 15 대구대학교 심리학과
전화: 053) 850-6366, Fax: 053) 850-6369, E-mail: gooya@daegu.ac.kr

여러 조직에서 수집된 자료를 통합하여 안정된 가중치를 산출하는 것이 하나의 방법이지만 이는 한 조직을 대상으로 한 검사개발도 어려운 국내 현실에서 실현가능한 대안이 아니다. 만약 전기자료가 유사 직종간 그리고 다른 조직간 안정된 타당도가 확보된다면 하나의 조직에서 개발된 전기자료를 타 조직에서도 이용할 수 있으며 규모가 작은 조직에서도 전기자료를 이용할 수 있는 계기가 될 수 있을 것이다.

전기자료는 배경자료(background data)라고도 하는데 이는 각 개인의 생활사와 관련된 사항(life history information)들에 관한 표준화된 형태의 문항 목록이다. 전기자료로 수행을 예측하고자 할 때 개인이 과거에 어떤 환경에서 무엇을 했느냐가 그 사람의 미래 행동을 잘 예측해 주리라고 전제한다(Owens, 1976). 이 전제는 이론적으로 사람들이 시간에 걸쳐 대체로 일관된 행동을 보일 것이라고 가정하는 행동 일관성 원리(behavioral consistency principle)에 바탕을 두고 있다(Dean, 1999). 행동 일관성 원리에서는 과거의 특정 상황에서 보인 행동 패턴은 미래의 유사한 상황에서도 그대로 나타날 가능성이 높을 것이라고 기대한다.

기본적으로 전기자료 검사는 행동 일관성 원리를 따르고는 있으나 미래의 행동과 동일한 과거 수행 측정치를 가지고 미래 행동을 예측코자 하는 것은 아니다. 왜냐하면 이전에 직장생활 경험이 없거나 상이한 직종으로 이직을 하는 사람의 경우 앞으로 근무하게 될 업무 상황을 과거에 경험한 바 없어 과거의 경험을 이용한 미래의 행동 예측은 어렵게 되기 때문이다. 이에 대해 Wernimont와 Campbell(1968)은 준거 수행영역의 행동과 동일하지는 않으나 인과적으로 관련되기

나 상관성이 높을 것으로 가정된 영역으로부터의 행동들까지 전기자료 문항에 포함시킬 것을 제안하였다.

1922년 최초로 영업사원의 수행을 예측하기 위하여 전기자료를 사용한 연구가 발표된 이후 현재까지 전기자료는 영업 수행의 예측뿐만 아니라 종업원의 퇴직이나 이직 및 사무직의 수행 예측, 그리고 연구 능력과 창의성과 전기자료와의 관련성 등 여러 직무 분야에서 다양하게 이용되고 있다(Childs & Klimoski, 1986). 이들 연구들에서 전기자료와 다양한 준거들간에 .30에서 .40 사이의 비교적 높은 타당도가 보고되고 있다(Hunter & Hunter, 1984; Nickels, 1994a, 1994b). 여러 직종 중에서 영업직의 수행 예측에서 전기자료가 가장 많이 활용되고 또 높은 타당도가 보고되고 있다. 국내 전기자료 연구들(강혜련, 1996; 박동건, 전인식, 2001; 이종구, 2001, 2002; 이종구, 윤창영, 임대열, 1999)도 주로 영업직을 대상으로 하였다. 전기자료가 영업직 수행예측에 많이 이용되는 주된 이유는 다른 직종에 비해 준거가 명확하다는 점 외에 영업직의 업무수행에서 요구되는 다음과 같은 특성들 때문이라고 볼 수 있다.

영업사원의 업무 수행에서 가장 두드러지게 요구되는 특성은 자율성(autonomy)과 거절(rejection) 경험에 대한 대처이다(Vinchur, Schippmann, Switzer, III, & Roth, 1998). 우선 영업사원은 집과 회사 밖에서 감독자 없이 업무를 수행한다. 따라서 상당한 자율권이 있으며 자신의 직관력과 설득력에 의지하여 출세해서 일해야 한다. 상황 압력에 의한 행동 규제가 적은 이러한 상황적 여건은 자신의 과거 경험의 표출을 용이하게 한다. 둘째, 다른 직종과는 달리 영업직은 대인 접촉 기회가 많고 수도 없이 거절당해 보기 때문

에, 이에 대한 감내력과 성취를 위한 지속적인 노력이 요구된다. 이러한 영업직의 특성은 일반 인지 능력보다 동기, 성격 특성 및 전기 자료가 영업수행을 더 잘 예측하게 해주는 원인이 될 수 있다. Vinchur 등(1998)의 메타분석 결과에서도 성격 5요인 중 성실성(conscientiousness)의 하위 요인인 성취지향성(achievement-orientation)과 전기자료가 영업사원의 수행을 가장 잘 예측하는 것으로 나타났다. 셋째, 전기자료는 점수화 방법의 특성상 대규모 자료가 요구되는데 특정 회사에서 이를 충족시켜 줄 만한 직무는 영업직이나 생산직 정도라는 것도 전기자료가 영업직에 많이 이용되는 이유 중 하나이다.

전기자료 타당도의 불안정성

전기자료 문항은 4가지 척도화 방법 - 개념적 척도화(rational scaling), 경험적 점수화(empirical keying), 요인척도화(factorial scaling), 및 하위집단화(subgrouping) - 에 의해 점수화될 수 있다.¹⁾ 4가지 척도화 방법 중 가장 보편적으로 오랫동안 사용되어 왔고, 가장 실제적이고, 높은 타당도를 확보할 수 있는 방법은 경험적 점수화 방법이다. 그러나 이 방법은 이론에 근거하지 않아 준거와 예측변수간의 관계에 대한 이해가 어렵고, 표본의 변화나 시간 경과에 따라 타당도가 다소 불안정하다는 문제점을 안고 있다. 또한, 경험적 점수화 방법은 외적 준거의 변화, 집단이 달라짐 또는 조직 내외적 환경이 변화함에 따른 타당도 감소가 점수화 방법 중 가장 큰 것으로 보고되고 있다

(이종구, 2001; 이종구, 윤창영, 임대열, 1999; Mitchell & Klimoski, 1982, 1986).

식품매장 판매사원 300명을 대상으로 한 이종구 등(1999)의 연구에서 여러 가지 점수화 방법 중 경험적 점수화 방법의 타당도가 가장 높게 나타났으나 교차타당화에서 가장 큰 타당도 감소현상이 나타났다. 그러나 이들의 연구는 문항 및 가중치 개발 집단과 교차타당화 집단의 표본수가 각각 194명과 104명으로 경험적 점수화 방법을 이용하기에는 지나치게 표본의 수가 적다는 문제를 지니고 있다. 한편 박동건과 전인식(2001)은 보험판매사원(생활설계사) 4813명을 대상으로 검사를 실시하고 이 중에서 대표본(4118명), 중표본(835명), 소표본(303명)을 뽑아 이들 자료를 각각 7대 3의 비율로 개발집단과 교차타당화 집단으로 분할한 후 타당도 변화를 알아보았다. 이들의 연구 결과, 경험적 점수화 방법을 이용했을 때 평균적으로 가장 높은 타당도를 보였고 타당도 감소의 폭도 가장 크게 나타났으나 감소된 후의 타당도도 다른 가중치 체계보다 높게 나타나 한번의 교차타당화 결과만 놓고 본다면 경험적 점수화가 가장 효율성이 높은 것으로 보인다. 또 표본의 크기가 클수록 경험적 점수화의 타당도 감소가 적게 일어났는데, 이들의 연구에서 대표본일 경우 개발집단과 교차타당화 집단의 타당도가 각각 .400과 .348로 상당히 안정적으로 나타났다. 이와 관련하여 9017명을 대상으로 경험적 점수화 방법을 이용하여 전기자료 검사를 개발한 Carlson, Scullen, Schmidt, Rothstein과 Erwin(1999)의 연구 결과, 전기자료의 타당도의 안정성이 매우 높게 나타났음은 물론 한 조직의 관리직을 대상으

1) 4가지 점수화 방법에 대한 자세한 내용은 강혜련(1996), 박영석(1995) 및 이종구 등(1999)의 논문에 제시되어 있다.

로 개발된 전기자료 검사를 타 조직의 관리직에도 그대로 이용할 수 있음을 보여주었다. 이들의 연구에서 하나의 조직에서 개발되고 가중치가 부여된 전기자료 검사를 24개 조직의 7334명의 관리자 표본에 대해 교차타당화를 시도한 결과 타당도의 평균이 .53이며 표준편차가 .05로 여러 조직에 걸쳐 타당도 차이가 크지 않은 것으로 나타나 전기자료의 일반화 가능성을 높였다. 또 여러 인구 통계적 특성이 다른 집단들간에도 타당도 차이가 크지 않은 것으로 나타났다.

따라서 대표본을 이용할 경우 경험적 점수화의 두 가지 문제 중 타당도 감소와 관련된 문제는 표본의 수를 크게 함으로써 어느 정도 해결되는 것으로 보인다. 그러나 연구 대상으로 삼은 조직의 규모가 작을 경우 적은 표본수로 인한 타당도 감소 문제는 여전히 나타날 것이다. 이때는 대개 여러 회사를 대상으로 대표본을 추출하여 제작하는 것이 권장사항이기는 하지만(Gandy, Dye, & MacLane, 1994) 국내에서 이런 형태로 검사를 제작하기란 어려운 실정이다.

요약하면, 전기자료의 점수화 방법 중 가장 바람직한 방법은 대표본을 대상으로 한 경험적 점수화 방법이다. 또 경험적 점수화 중에서도 전기자료 문항의 특성이 대개 명명 또는 서열 척도 수준에서 제작되기 때문에 회귀나 등가중 방법보다 WAB 가중 방법이 보다 적절할 수 있다.

경험적 점수화 방법에는 준거와 예측 변수인 전기자료 문항의 수량적 특성에 따라 여러 가지 가중치 부여 방법이 가능한데, 크게 두 가지로 대별된다(Hogan, 1994). 그 하나는 만약 각 문항에 5가지 선택지가 있다면 각 선택지 하나하나에 가중치를 부여하는 선택지 단위 점수화 방법(option-keying method)으로 Strong의 WAB 기

법이 대표적인 예다(Guion, 1965). 이 기법은 특정 수행 준거 상에서 우수-부진 집단간 통계적으로 유의한 차이를 보이는 문항을 대상으로 두 집단간 각 선택지에 응답한 비율 차이를 토대로 가중치를 부여하고 문항들의 가중치들을 합산하여 이 합산된 점수가 준거를 얼마나 잘 예측해 주는지를 따지는 기법이다. 이 기법을 적용하는 주된 상황은 준거가 양분집단(우수-부진)이고 표본의 수가 큰 상황이다.

다른 하나는 문항의 반응 점수를 이용하는 문항단위 점수화 방법(item-keying method)인데 수량적 척도에 적용되며 이는 다시 문항별로 중요도에 따라 다른 가중치를 부여하는 경우와 등가중치(equal or unit weighting)를 부여하는 경우가 있다. 문항별로 차별적인 가중치를 부여하는 방법 중 가장 대표적인 것은 회귀계수를 가중치로 이용하는 방법이다. 이중 어떤 가중치 체계를 선택하느냐는 예측력(predictive power)과 가중치 이용의 편의성에 달려있다. 대수적 정보통합 방법과 예측의 정확도간의 비교연구 결과들(Dawes, 1971, 1982; Goldberg, 1968, 1971)에서 다소 복잡한 회귀 가중치와 보다 단순한 등가중(equal weighting) 방법이 다른 가중치 부여 방법에 비해 예측력이 높고 또 이 둘간에는 예측력 차이가 보이지 않는 것으로 일관되게 보고되었다.

앞에서 살펴보았듯이, 경험적 점수화에서 점수부여 체계를 개발하는 데는 개발집단의 사례수가 대단히 중요하다(Hogan, 1994). 충분한 사례수가 확보되어야 수행에서의 우연적 차이를 제거하고 준거집단의 수행차이를 정확하게 변별해 줄 수 있다. 우연에 의해 오는 예측력의 과장된 증가는 예측 변수의 수에 따라 증가하고 사례수가 증가함에 따라 감소하기 때문에 예측변수대비 사례

수의 비율을 최대화 시켜줄수록 안정적인 수행에 측을 가능하게 한다.

Nunnally(1978)는 사례수와 문항수의 최소 비를 5 대 1 또는 10 대 1 정도 되기를 권장하고 있는데 이를 기준으로 한다면 100문항을 이용할 경우 500에서 1000명 정도의 사례수를 요구한다고 보면 된다. 그러나 이 기준은 문항단위 가중치 부여방식 (item-keying) 방법을 이용할 때의 기준이고 선택지단위 가중치 부여방식 (option-keying)을 이용할 경우(예, WAB 가중) 개별 문항 단위로 가중치가 유도되는 것이 아니라 개별 문항의 반응선택지마다 가중치가 유도되기 때문에 이때 총 예측변수의 수는 모든 문항의 반응선택지의 수와 같아진다. 예컨대 문항의 수가 100개이고 각 문항의 선택지가 3개이면 고려해야 하는 총 예측변수의 수는 300개인 셈이 되어 Nunnally가 제안한 사례수 기준을 충족시키려면 1500명에서 3000명의 사례수가 필요하게 된다. 국내 단일 회사에서 이 정도의 사례수 확보가 가능한 회사는 많지 않다.

이종구(2002)의 연구에서는 사례수의 문제를 부분적으로 해결하기 위해 한번의 반응을 토대로 다른 반응을 예측하기보다 동일인으로부터 여러 번의 반응을 통합하여 다른 반응을 예측하는 것이 예측 오차를 줄일 가능성이 높다고 보고 2회 이상의 반응자료를 통합하여 가중치를 계산하여 타당도의 안정성을 검토하였다. 그 결과 한번의 반응자료로 채점기준을 설정하는 것보다 동일인이라고 할지라도 2회의 반응자료로 채점기준을 마련하는 것이 교차타당화시 타당도 감소가 줄어드는 것으로 나타났다. 즉 1차와 2차 자료를 단순히 통합한 자료를 대상으로 가중치를 유도해낸 후 이 가중치를 1, 2, 3차 및 교차타당화 자료에

적용하여 준거와의 상관을 구했을 때 타당도가 안정적임을 보였다.

그러나 2회의 반응을 통합해도 전기자료 가중치를 산출할 만한 자료가 되지 않는 회사의 경우 회사자체에서 전기자료 검사를 개발하기는 어렵다. 단, 동일 직종 또는 유사 직종의 타 회사에서 개발된 가중치(또는 scoring key)를 그대로 이용했을 때 안정적인 타당도가 확보된다면 소규모 회사에서도 타 회사에서 개발된 문항과 가중치를 그대로 이용할 수 있을 것이다.

본 연구에서는 이의 가능성을 확인하기 위해 여성 보험판매사원을 대상으로 한 이종구(2001)의 연구와 여성 텔레마케터를 대상으로 한 이종구(2002)의 연구에서 이용된 자료를 토대로 각각의 자료에서 가중치를 산출한 후 나머지 자료를 교차타당화 자료로 삼아 안정적인 타당도가 유지되는지를 확인해보고자 한다.

방 법

1. 문항구성

여성 보험 판매사원을 대상으로 실시한 전기자료 문항은 성장기 경험과 최근 경험으로 나누어 다음의 원천으로부터 구성되었다. 먼저 성장기 경험을 측정하는 문항은 'MBC 다큐멘터리 성공시대'에 등장한 인물 중 영업경험이 있는 기업인(26인) 자료에 대한 내용분석에서 유도된 396문항과 Owens(1976)의 전기자료 검사 114문항을 합친 510 문항 중에서 선정하였다. 510문항 중 영업수행에 영향을 주는 것으로 알려진 심리적 구성개념(Vinchur, et al., 1998)(활동성, 낙천성, 성실

성, 혁신성, 독창성)과 유의한 상관을 보인 문항 중 상관계수가 큰 순서로 60문항이 선정되었다. 최근 경험을 측정하기 위한 문항은 우수 보험판매사원을 대상으로 한 면접과 그리고 LIMRA career profile의 200문항 중 영업수행과 관련된 26문항이다. 이 전기자료 검사는 총 86문항으로 이루어져 있다.

여성 텔레마케터를 대상으로 실시한 전기자료 검사는 우수 텔레마케터 면접을 통한 문항수집 절차를 제외하고는 여성 보험판매사원용 문항개발 절차와 동일하였다. 단, 검사실시시간에서 다소 여유가 있어 보다 느슨한 기준에 의해 문항을 선정하여 총 219문항으로 이루어진 검사를 이용하였다. 문항들은 성장기 경험을 측정하는 117문항, 입사전 경험 102문항들이다. 본 연구에서 분석대상으로 삼은 문항은 여성 보험판매사원을 대상으로 실시된 86문항과 여성 텔레마케터를 대상으로 실시된 219문항 중 공통으로 실시된 75문항이다.

2. 자료수집 및 조사대상자

국내 S 보험사의 여성 보험판매사원 1693명에 대한 자료는 2001년 상반기에, 국내 L사의 417명의 텔레마케터를 대상으로 한 자료는 2001년 12월에서 2002년 2월까지 걸쳐 수집되었다.

월평균 순소득을 토대로 수행이 중간이상인 보험판매사원 중에서 무선적으로 선정하였다. 전

국 91개 지점에 각 30부씩 총 2730부를 배포하여 1726부(63.22%)가 회수되었다. 이중 준거 자료가 누락된 33명을 제외한 1693명의 자료가 시초분석 대상 자료였다. 최종 분석 대상집단은 1693명의 자료를 준거수행을 기준으로 상(우수), 중, 하(부진) 집단으로 3등분하여 중간집단을 제외한 상하 집단 1130명이다. 집단의 구분은 재직기간별로 이루어졌다. 재직기간이 길수록 보험판매량이 증가하는 추세를 보였고, 재직기간이 7년 이상일 경우 판매량 증가를 보이지 않았다. 따라서 재직기간을 1년에서 7년 이상까지 1년 단위로 집단을 구분하여 각 재직기간내에서 준거를 기준으로 상, 중, 하 집단으로 구분하였다. 표 1에서 우수집단과 부진집단이 각각 566명(50.09%)과 564명(49.91%)으로 비율이 거의 유사하다. 분석 대상집단 1130명의 평균연령은 39.53세이고 표준편차는 7.24였다.

텔레마케터들을 대상으로 한 검사실시는 근무시간 중에 통제된 방에서 10-20명씩 집단적으로 이루어졌다.

최종분석 자료는 업적 평가를 기준으로 자료를 3등분하여 중간집단을 제외한 상하집단이다. 업적평가는 L사에서 제공한 인사고과 점수인대 개인별 수주, 클레임 수, 고객불만처리 등의 점수를 통합한 점수이다. 우수집단의 최하위 및 부진집단의 최상위 동점자료는 모두 우수 또는 부진 집단에 포함시켰다. 분석자료에 포함된 자료의 수는 282명으로 우수집단이 140명(49.65%)이었으며

표 1. 조사대상자 수

	우수	부진	전체
보험판매사원	566(50.09)	564(49.91)	1130(100.00)
텔레마케터	140(49.65)	142(50.35)	282(100.00)
전체	706(50.00)	706(50.00)	1412(100.00)

부진집단이 142명(50.35%)이었다(표 1 참조). 282명의 평균 연령은 27.15세(SD=3.39)였으며 평균 재직기간은 1.08년(SD=.79)이었다.

분석과정에서 한번은 보험판매사원 집단을 개발집단으로, 텔레마케터 집단을 검증(또는 교차타당화) 집단으로 이용하였고, 다른 한번은 역으로 하였다.

결 과

1. 문항선별

문항선별은 1130명의 보험판매사원을 대상으로, 그리고 282명의 텔레마케터 집단을 대상으로 각각 이루어졌다. 두 집단에 공통적으로 사용된 75문항 각각에 대해 우수-부진집단×문항선택지의 빈도자료를 대상으로 한 χ^2 검증 결과를 토대로 하였다. χ^2 검증 결과 .05 수준에서 유의한 문항을 최종적으로 선별하였다. 보험판매사원 집단에서는 37문항이, 텔레마케터 집단에서는 33문항이 최종 선별되었다. 각 집단별로 선별된 문항은 표 2에 제시되어 있다.

표 2 선별된 문항

문 항	보험 판매	텔레 마케터
입사전 사내에 사람 아는 정도	*	-
입사전 가족생계 책임정도	-	*
입사전 가족의 재산	*	*
입사전 가장 오래 일한 직종	*	*
입사전 주당 집과 직장이외의 활동시간	*	-
입사전 이틀부를 정도 친한 이웃의 수	-	*
입사전 주당 친척/친구의 전화/방문수	*	*
입사전 1년간 소득	*	*

문 항	보험 판매	텔레 마케터
입사전 월평균 최소 생활비	*	-
입사후 첫해 기대소득	*	-
입사후 3년차 기대소득	*	-
입사전 4년간 정기적 운동횟수	-	*
입사전 직장의 수	*	-
입사전 현장방문 직업경험	*	-
중고때 쉬는 시간에 주변에 친구가 모인 정도	*	*
학창시절 리더 경험	-	*
초등학교 학급성적	-	*
남들이 비웃어도 할 이야기는 끝까지 한다	*	*
함께 있을 때 가장 말을 많이 하는 편	*	*
잡기에 능한 정도	-	*
중고교때 외모에 신경을 많이 씀	-	*
초등학교 때 동화를 많이 읽음	-	*
중고교 때 남의 발표시 답답	*	*
자기 싫어하는 친구에게는 저준다	*	*
남들이 부러워하는 특기가 있다	-	*
이 길이 아니다 싶으면 빨리 중단한다	*	*
초중학교 이웃 어른들로부터 귀여움	*	*
일을 끝내기 위해 밤을 많이 새는 편	*	-
물건 어디 두었는지 자주 못찾음	*	-
결심한 것 반드시 지킴	*	*
유행 헤어스타일과 의상을 즐겨 입음	*	*
고교때 등수목표 정해놓고 공부	*	*
나한테 물어봐(내가 최고) 하는 것 있음	*	-
중고교때 시험보고나면 곧바로 채점	-	*
친구보다 낮은 성적받으면 다음에 이기려	*	*
고교때 똑똑하다는 말을 가장 듣고 싶었다	*	*
집보다는 밖에서 더 인정	*	*
중고교때 미술/음악대회에서 수상경험정도	*	*
중고교 등산/놀이갈 때 가장 많이 준비	*	-
중고교때 모임에서 친구들을 이끈다	*	*
고교때 토론을 좋아함	*	-
학창시절에 주로 어울린 연령대	*	*
학창시절에 웅변대회 참가정도	*	-
주변사람들의 감정에 민감	-	*
당신의 친척들의 생활 수준	*	-
목표설정 정도	*	*
고교때 한과목이라도 일등목표로 공부	*	*
운동(탁구/테니스/볼링/수영)빈도(최근 4년)	*	-

* $p < .05$ - ns

2. 타당도 비교

우선 보험판매사원 집단을 개발집단으로 했을 때의 타당도 계산과정을 설명하면 다음과 같다. 준거는 월평균 소득이다. 보험판매사원의 영업실적은 연차별로 누적되고 또 소득이 낮은 사람은 퇴직하는 경우가 많기 때문에 재직기간과 소득간에는 높은 상관관계를 지닐 수 있다. 따라서 전기자료의 타당도 계산 시 재직기간과 준거간의 관련성이 배제되어야(partial out) 한다. 이를 위해 본 연구에서는 월평균소득을 종속변수로, 재직기간을 독립변수로 하여 단순 회귀분석을 실시한 후 종속변수 값에서 예측된 종속변수의 값(재직기간효과)을 뺀 나머지 값(예측오차)을 준거로 삼아 타당도를 계산하였다.

전기자료의 가중치 부여는 WAB 가중방식에 의해 이루어졌다. WAB 방식에 의한 가중치 부여는 선별된 37개 문항을 대상으로 Strong의 가중치 계산 방법에 의거하여 계산하였다. WAB 방식(Strong의 가중치 계산방법)에서는 우수집단과 부진집단의 문항별 응답비율을 계산하여 두 집단의 응답 비율의 차이를 계산한 후 비율차이에 따라 선택지마다 다른 가중치를 부여한다(자세한 내용은 1965년에 출간된 Guion의 Personnel Testing의 388쪽 참조). WAB 점수는 선별된 37 문항에서 받은 가중치들의 합이며, 이 합산 점수가 타당도 계산에서 예측변수로 이용된다. 개발집단을 대상으로 구한 문항별 가중치는 교차타당화 집단에도 그대로 적용되었다.

텔레마케터 집단을 가중치개발 집단으로 이용했을 때의 가중치 산출도 WAB 가중 방식을 이용하였다. 33개 문항에 대해 업적평가를 기준으로 나는 우수-부진 집단간의 문항 반응 비율차이를

이용하여 문항들의 가중치를 계산하고 이 가중치는 교차타당화 집단에도 그대로 적용하였다.

표 3은 각 집단을 가중치개발 집단으로 이용했을 때 개발집단과 교차타당화 집단의 타당도 분석 결과이다. 개발집단이 보험판매사원 집단(1130명)이면 교차타당화 집단은 텔레마케터 집단(282명)이고, 개발집단이 텔레마케터 집단이면 교차타당화 집단은 보험판매사원 집단이 된다.

표 3. 타당도 분석 결과

가중치 개발집단	타당도 계수	교차타당도 계수
보험판매사원 (1130명)	.392*	.294*
텔레마케터 (282명)	.597*	.254*

* $p < .01$

표 3의 타당도 분석 결과 대표본인 보험판매사원 집단을 개발집단으로 했을 때의 타당도는 .392, 교차타당화시(텔레마케터 집단)의 타당도는 .294로 나타나 타당도 감소가 크게 일어나지는 않았다. 반면, 소표본인 텔레마케터 집단을 개발집단으로 이용했을 때 개발집단의 타당도는 .597로 매우 높게 나타났으나 교차타당화시의 타당도는 .254로 유의하기는 하지만 상당한 타당도 감소가 일어났다. 물론 이 결과만으로 집단 특성의 차이는 논하기 어렵다. 표본의 크기만을 놓고 본다면 가중치 개발집단의 자료로 대표본을 이용했을 때 교차타당화 집단의 타당도가 안정되는 것으로 보이며 이 결과는 박동건과 전인식(2001)의 연구 결과와도 일치한다.

논 의

하나의 조직에서 유도된 전기자료 문항의 가중치를 유사 직종의 타 조직에 적용했을 때 안정적인 타당도를 얻을 수 있는지를 확인하고자 시도된 본 연구의 결과를 요약, 논의하면 다음과 같다.

본 연구 결과 대표본을 가중치 개발집단으로 이용하고 소표본을 교차타당화 집단으로 이용했을 때나 그 역으로 자료를 이용했을 때 모두 가중치 개발집단의 타당도는 물론 교차타당화 집단의 타당도가 유의하게 나타나 한 조직에서 개발된 전기자료의 가중치가 일반화 가능함을 보여주었다. 그러나 개발집단의 타당도는 소표본(282명의 텔레마케터, $r=.597$) 일 때가 대표본(1130명의 보험판매사원, $r=.392$)일 때보다 높게 나타났으나 개발집단으로 대표본을 이용했을 때가 소표본을 사용했을 때보다 교차타당화 집단의 타당도 감소가 적음과 동시에 타당도도 더 높게 나타났다. 이 결과는 대규모 조직에서 개발된 전기자료 검사는 소규모 조직의 유사직종에 적용할 수 있음을 시사해 준다.

지금까지 전기자료의 타당화 연구에서 가장 많이 사용되어온 문항점수 부여방법은 경험적 점수화(empirical keying) 방법이다. 이 방법은 사용하기 쉽고 높은 타당도를 얻을 수 있다는 장점이 있으나 타당도가 매우 불안정하고, 타당도에 대한 설명적 근거, 즉 문항-준거간 이론적 연결고리를 제공해주기 어렵다는 문제가 계속 지적되어 왔다. 또 타당도를 높이기 위해 WAB 기법을 비롯하여 다양한 가중치부여 방법이 제안되어 타당도의 개선이라는 목표는 어느 정도 달성되었으나 그 어느 기법도 표본의 변화나 시간경과에 따른

타당도의 불안정성을 해결하고 있지는 못하다.

또한 조직내 직무간, 유사 직무의 조직간 일반화 연구가 드물어 다른 조직에서 개발된 전기자료 검사 목록을 타 조직에 그대로 이용하는 사례가 드물다. 물론 타 조직에서 개발된 문항목록을 그대로 이용할 수는 있으나 문항별 가중치 체계는 해당 조직에서 개발해야 하는데 조직의 규모가 작을 경우 불안정한 가중치가 유도되어 이로 인해 교차타당도가 매우 낮게 보고되고 있다.

본 연구에서는 전기자료 문항-수행준거간 이론적 연결고리를 찾는데 일조하고 보다 안정된 타당도를 얻기 위해 다양한 원천으로부터 597문항(성장기 경험 397문항, 입사전 최근경험 200문항)을 수집하였다. 이들 문항들 중에서 타당화 연구에 사용한 문항들에 대한 용이한 해석을 위해, 그리고 타당도를 높이기 위해 선별기준으로 수행에 영향을 주는 것으로 알려진 심리적 구성개념들을 이용하였다. 그 결과 개발집단의 타당도가 .39와 .60(표 3 참조)이었으며 교차타당화 집단의 타당도가 각각 .29와 .25로 나타나 Hunter와 Hunter(1984)가 보고한 전기자료의 타당도 .30~.40보다 다소 높거나 유사하게 나타났다.

본 연구에서 보험판매사원 집단을 대상으로 가중치 개발에 사용한 문항은 37문항이며 텔레마케터 집단을 대상으로 가중치 개발에 사용한 문항은 33문항이다. 가중치 산출은 문항단위 가중치 부여 방식(item-keying)이 아닌 문항내 선택지단위 가중치 부여방식(option-keying)을 택했으며 각 문항별 평균적인 선택지의 수는 5개였으므로 문항들에 포함된 선택지의 수를 문항수로 간주한다면 37문항의 경우 185문항, 33문항의 경우 165문항인 셈이 된다. 가중치 산출에 필요한 최소사례수에 대한 Nunnally(1978)의 권고에 따르면 각

각 925명과 825명이 된다. 보험판매사원 집단 1130명은 Nunnally의 권고 사례수인 925명을 초과하지만 텔레마케터 집단 282명은 권고사례수인 825명에 훨씬 못 미친다. 본 연구 결과에서 텔레마케터 집단을 가중치 개발 집단으로 이용했을 때 타당도 감소가 크게 일어난 이유도 적은 사례 수 탓으로 볼 수도 있을 것이다.

본 연구의 결과가 폭넓게 수용되려면 몇 가지 해석상의 주의와 이론적인 검토 및 체계적인 보완 연구를 필요로 한다. 첫째, 두 집단에 공통적으로 실시된 문항은 75문항이었으며 본 연구에서는 이들 문항만을 분석대상으로 삼았다. 그러나 실제 실시된 문항은 보험판매사원용은 86문항, 텔레마케터는 219문항으로 구성된 검사목록이었다. 따라서 75개의 공통문항 외에 상이한 검사문항이 공통문항에 대한 반응편향을 유도했을 가능성을 배제하기 어렵다. 단 상이한 문항목록을 구성하고 실시한 이유는 각 회사의 검사실시 여건이 달랐기 때문이며 각 직무 수행 예측에 특유한 문항이 포함되었기 때문이다.

둘째, 앞서 논의된 교차타당화 집단의 타당도 감소문제와 함께 개발집단의 타당도 크기가 매우 다르게 나타난 점에 대해서도 주목할 필요가 있다. 가중치 개발 집단의 타당도는 텔레마케터의 경우 .60으로 나타났으나 보험판매사원의 경우 .39로 나타나 상당한 차이를 보였다. 텔레마케터 집단의 높은 타당도가 소표본으로 인한 표본 변동 탓도 있으나 보험판매사원 집단에 비해 비교적 동질적인 업무환경 탓에 수행예측이 더 용이했기 때문일 수도 있다. 한편 보험판매사원의 타당도 .39가 텔레마케터보다 더 낮게 나온 결과는 보험판매사원용 검사제작시 검사시간의 제한으로 인해 검사길이에 상당한 제약을 받았기 때문에

문항 선별 과정에서 영업사원의 핵심역량(competency)을 측정해주는 문항들 중 다수가 제외되었기 때문일 수 있다(이종구, 2001). 셋째, 이종구(2002)의 연구에서 2차례에 걸친 응답결과를 통합하여 표본의 크기를 늘린 연후에 가중치를 유도한 결과 비교적 안정된 타당도를 유지하는 것으로 나타났다. 이와 동일한 방법은 아니지만 대규모 조직에서 개발된 가중치를 소규모 조직에서 이용할 경우 자료를 지속적으로 통합해가면서 가중치를 조정하는 것도 타당도를 높이면서 타당도의 안정성을 높이는 한가지 방법일 수 있다. 추후 전기자료에 대한 보다 다양한 직종간 일반화 연구가 이루어진다면 가중치 개발 집단과 유사한 소규모 조직은 물론 다른 직종에 대해서도 전기자료의 이용이 확대될 수 있을 것이다.

참고문헌

- 강혜련(1996). 경험적 척도화에 의한 전기자료의 타당화 연구. 한국심리학회지: 산업 및 조직, 9(2), 1-16.
- 박동건, 전인식(2001). 전기자료(Biodata) 문항의 가중치 부여체계간의 타당도 연구: 분석집단크기에 따른 비교연구. 한국심리학회지: 산업 및 조직, 14(1), 101-114.
- 박영석(1995). 이직자의 생활사. 한국심리학회지: 산업 및 조직, 8(1), 75-102.
- 이종구(2001). 전기자료(Biodata)의 타당화 연구: 보험판매사원의 수행 예측. 한국심리학회지: 산업 및 조직, 14(3), 83-107.
- 이종구(2002). 소표본에서 오는 전기자료 타당도의 불안정성 개선 방법 모색. 한국심리학

- 회지: 산업 및 조직, 15(3), 73-87.
- 이종구, 윤창영, 임대열(1999). 영업수행의 예측에서 전기자료의 이용. 한국심리학회지: 산업 및 조직, 12(1), 141-153.
- Carlson, K. P., Scullen, S. E., Schmidt, F. L., Rothstein, H., & Erwin, F. (1999). Generalizable biographical data validity can be achieved without multi-organizational development and keying. *Personnel Psychology, 52*(3).
- Childs, A. & Klimoski, R.(1986). Successfully predicting career success: An application of the biographical inventory. *Journal of Applied Psychology, 71*, 3-8.
- Dawes, R. M.(1971). A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist, 34*, 571-582.
- Dawes, R. M.(1982). The robust beauty of improper linear models in decision making. In D. Kahneman, P. Slovic, & A. Tversky(Eds.), *Judgment under uncertainty: Heuristics and biases*(pp. 391-407). Cambridge, UK: Cambridge University Press.
- Dean, M. A. (1999). *On biodata construct validity criterion-related validity, and adverse effect*. Unpublished doctoral dissertation, Louisiana State University.
- Gandy, J. A., Dye, D. A., & MacLane, C. N. (1994). Federal government selection. In Stokes, G. S., Mumford, M. D., & Owens, W. A.(Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. Palo Alto, CA: CPP books.
- Goldberg, L.R.(1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist, 23*, 483-496.
- Goldberg, L.R.(1971). Five models of clinical judgment: An empirical comparison between linear and nonlinear representations of the human inference process. *Organizational Behavior and Human Performance, 6*, 458-479.
- Guion, R. M. (1965). *Personnel testing(ch.13)*. McGraw-Hill.
- Hogan, J. B.(1994). Empirical keying of background data measures. In Stokes, G. S., Mumford, M. D., & Owens, W. A.(Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. Palo Alto, CA: CPP books.
- Hunter, J. E., & Hunter, K. F.(1984). Validity and utility of alterntive predictors of job performance. *Psychological Bulletin, 96*, 72-98
- 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.
- Mitchell, T. W., & Klimoski, R. J. (1986).

- Estimating the validity of cross-validity estimation. *Journal of Applied Psychology*, 71. 311-317.
- Nickels, B. J.(1994a). Generation procedures for biodata items. In Stokes, G. S., Mumford, M. D., & Owens, W. A.(Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. Palo Alto, CA: CPP books.
- Nickels, B. J.(1994b). The nature of biodata. In Stokes, G. S., Mumford, M. D., & Owens, W. A.(Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction*. Palo Alto, CA: CPP books.
- Nunnally, J. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Owens, W. A.(1976). Background data. In M.D. Dunnette(Ed.), *Handbook of industrial and organizational psychology*. Chicago: Rand McNally.
- Vinchur, A.J., Schippmann, J.S., Switzer, III, F.S., & Roth, P.L.(1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology*, Vol. 83, No. 4, 586-597.]
- Wernimont, P. F., & Campbell, J. P. (1968). Signs sample, and criteria. *Journal of Applied Psychology*, 52, 372-376.

한국심리학회지 : 여성
The Korean Journal of Woman Psychology
2002. Vol. 7, No. 3, 73 - 86

Validity Generalization Study of Biographical Data for Two Female Sales Job Groups

Lee, Jong-Goo
Daegu University

Kim, Myoung-So
Hoseo University

Youn, Chang-Young
Daegu University

This study was conducted to verify the stability of validity in two female sales jobs(1130 life insurance sales women and 282 female tele-marketers). After a series of item analyses were conducted for each group, 37 items and 33 items among 75 biographical items were included to the validation study, respectively. The resulting validities and cross-validities were generalizable, regardless of whether sample size of item-keying group was large or small. But, large sample-based empirical keys showed larger cross-validity than small sample-based empirical keys. The finding in this study show that large sample-based keying of biodata scale can yield generalizable validities. And, the current findings suggest that consortium-based, multiple-organization biodata researches are needed.

Key words: biographical data, biodata, validity generalization, empirical keying, female sales job