

직무교육훈련 평가수준들간의 관계

김진모

LG 인화원

The Relationships between the Levels of Evaluation of the Training & Development for Job Skills

Jin Mo Kim

LG Academy

Summary

The purpose of this study was to analyze the relationships among the levels of training & development evaluation (reaction, learning, transfer). The study has been conducted on 730 trainees who attended in the basic accounting program in L training and development institution through three incidents of tracked research such as reaction survey right after the conclusion of training, learning evaluation through test, and an evaluation of the transferability after 3 months of training.

Questionnaires and test papers for analyses were used after their reliability, validity, difficulty, and discrimination have been verified on a pre-test. The research has been conducted for six months from 4 March 1996 to the end of August 1996, and data have been collected through direct research and survey through mail. The collected data have been worked on at SAS program for Windows with a statistical significance level of 5%. Statistical method that had been used was Pearson's correlation coefficient.

The result and conclusion acquired from this study were as follows:

Between reaction and learning, learning and transfer of training, only a weak positive correlation exists and explanation or prediction variance showing hierarchical relationship was quite weak with 1%. Thus, this research not only does not strongly support Kirkpatrick(1976)'s hierarchical model of reaction→learning→transfer, but also indicates that the separate measurement on each levels of training evaluation needs to be done. On the other hand, there was a relatively strong positive correlation between reaction and transfer of training.

Based on the result, the conclusion, and the restriction perceived through this study, the following suggestions were made.

1. There is a need to empirically analyze and verify the hierarchy of all levels of training evaluation including the evaluation of the fourth level (result) such as organizational productivity, organizational satisfaction, and separation rate.
2. A great deal of efforts will be needed to systematically analyze what the relationships are among the methods measuring the level of evaluation of the training and development, and to apply this result to the training field.

I. 서 론

1. 문제의 제기

일반적으로 기업의 교육훈련을 평가하는데 가장 널리 쓰이는 모델은 Kirkpatrick(1959, 1976, 1987)의 교육훈련 결과의 위계적 모델이다(Al-ammar, 1994:48). 이 모델에 따르면 교육훈련의 유효성은 다음의 기준(criteria)에 의하여 측정될 수 있다.

① 반응(reaction): '교육훈련 참가자들이 프로그램에 얼마나 만족했는가?' 와 같은 질문을 통해 교육훈련 프로그램에 대한 참가자들의 느낌과 선호도를 측정한다.

② 학습(learning): '교육훈련 참가자들이 어떤 지식과 스킬을 배웠는가?' 또는 '어떤 태도가 변화되었는가?' 와 같은 질문을 통해 원리, 사실, 기술 등에 대한 이해와 학습정도를 측정한다.

③ 행동(behavior): '교육훈련의 결과로 직무 행동상에 어떤 변화가 일어났는가?' 와 같은 질문을 통해 직무현장에서의 적용정도를 측정한다.

④ 결과(results): '교육훈련이 끝난 후, 질의 향상 및 양의 증가, 비용절감 등과 같은 측면에서 어떠한 유형(tangible)의 향상된 결과가 있었는가?' 와 같은 질문을 통해 프로그램의 조직에 대한 기여도를 측정한다.

각 평가수준별 교육훈련의 결과(outcome)는 위계성이 있어 다음 수준의 결과에 영향을 미친다. 즉, 교육훈련 참가자의 프로그램에 대한 만족도는 학습에 중요한 영향을 미치고, 교육훈련 프로그램의 내용은 직무현장에서의 행동 개선을 위해 어느 정도까지는 숙달되며, 행동 변화는 생산의 질이나 양과 같은 결과의 긍정적인 변화에 중요하게 작용한다는 것이다.

사실 많은 학자들간에는 Kirkpatrick이 주장하는 평가의 4단계가 위계적으로 선형적인 관계를 갖는다고들 가정하고 있다(Mathieu 등,

1992).

Hamblin(1974)은 교육훈련 참가자의 반응이 다른 평가수준과 정적 상관이 있고, 다른 평가수준의 선행조건이라고 보고 있다.

실제, 수 많은 교육훈련의 평가수준에 대한 연구결과는 교육훈련에 대한 만족과 학습, 행동변화가 함께 일어남을 증명함으로써 이와 같은 위계적 모델을 간접적으로 지지하고 있다(Fromkin, Brandt, King, Sherwood, & Fisher, 1975; Latham, Wexley, & Purcell, 1975).

Clement(1978)도 경로분석을 통해 교육훈련 참가자의 반응은 학습에 인과적 영향을 미치고, 학습은 행동변화에 중요한 영향을 미친다는 사실을 보여 줌으로써 이 위계적 모델을 강력하게 지지한 바 있다.

Trefz(1991) 또한 교육훈련에서 제공한 방식에 큰 만족을 보인 참가자들은 그렇지 않은 참가자들 보다 자신의 직무에 지식과 기술을 더 많이 활용하는 경향이 있음을 확인하였다.

Faerman과 Ban(1993)은 감독자 교육훈련 프로그램에 참가한 교육훈련생의 만족도와 교육훈련 후에 나타난 그들의 행동변화간에는 어느 정도 강한 관계가 있음을 밝혀 냈다. 그들은 몇몇 연구들에서 밝혀졌던 반응과 행동변화간의 약한 관계는 행동변화의 측정과 관련된 일종의 방법론적인 문제에 기인한 것이라고 주장하였다.

Al-ammar(1994)의 연구에서도 교육훈련 프로그램에 대한 참가자들의 반응이 교육훈련의 직무에 대한 전이와 유의미한 관계가 있었고, 이도형(1995)도 학습은 전이에 직접적인 영향을 미친다고 보았다.

반면에, 평가수준들간에 선형적 관계(linear relationship)가 있다고 보는 것은 타당하지 못하다고 주장하는 연구결과들도 있다. 특히 반응과 학습의 관계에서 그렇다. Alliger와 Janak(1989)는 Kirkpatrick의 네 가지 기준들간의 관계를 검증한 결과, 반응과 학습의 관계가 약하다는 사실을 확인했다. 사실, 그들은 이 두 기

준간의 부적인 관계는 오히려 논리적일 수 있다고 주장한다. 열심히 공부해야 되는 즉, 보다 많은 학습이 일어날 것으로 예상되는 과정들은 인기가 덜 할 수도 있다. 그러나 다른 기준들 즉, 학습, 행동의 변화, 그리고 결과에 대해서는 그 관계가 훨씬 더 강하게 나타났다.

Noe와 Schmitt(1986)도 네 가지 기준들간의 선형적 관계에 대해서는 일부만을 지지하는 결과를 제시했다. 그들의 연구에서는 반응과 학습, 학습과 행동변화는 서로 관련성이 없었고, 단지 행동변화 결과간에만 관계가 있는 것으로 확인되었다.

Mathieu 등(1992)은 반응과 학습이 왜 선형적인 관계가 있는지에 대한 이론적인 근거는 없을 지라도, 반응은 동기를 통해 학습과 관련이 있을 수 있다고 주장한다. 이 주장은 반응이 동기와 학습의 관계를 매개하는 것으로서 작용하고 있다고 확인한 그들의 연구에 의해 뒷받침되었다.

따라서, 이 연구에서는 교육훈련의 각 평가 수준들간에는 상호 관계성이 어느 정도 존재하는지를 알아보고자 한다. 만약 많은 학자들의 주장대로 평가수준들간에 상호 관계성이 있거나 강하다면, 실제 대부분의 평가가 만족정도 측정에 머물러 있는(권대봉, 1993:10) 우리의 현실을 감안할 때, 교육훈련 참가자의 반응 평가만으로도 학습이나 적용 등을 설명 혹은 예측할 수 있게 되므로 그 의의는 매우 크다고 볼 수 있다.

2. 연구목적

이 연구의 목적은 교육훈련의 각 평가수준들(반응, 학습, 전이)간의 관계를 구명하여 이론적 가정을 확인하는데 있다.

3. 연구문제와 가설

이 연구목적을 달성하기 위하여 다음과 같

이 연구문제를 설정하였다.

연구문제 : 교육훈련 참가자의 반응과 학습 간, 학습과 전이간, 반응과 전이간에는 서로 어떤 관계가 있을 것인가?

이 연구문제를 검증하기 위한 연구가설은 다음과 같다.

가설 1. 교육훈련 참가자의 반응은 학습과 정적인 상관관계가 있을 것이다.

가설 2. 교육훈련 참가자의 학습은 전이와 정적인 상관관계가 있을 것이다.

가설 3. 교육훈련 참가자의 반응은 전이와 정적인 상관관계가 있을 것이다.

4. 연구의 제한

1) 이 연구의 대상은 연구결과의 신뢰성과 조사의 용이성 등을 고려하여 편의적 추출방법을 통해 선정하였다. 즉, 국내 기업중 교육훈련에 대한 투자가 비교적 많고, 교육훈련의 양과 질적인 면에서도 수준급이라고 평가되는 특정 기업연수원을 대상으로 학습정도의 측정이 비교적 용이한 직무능력 향상 프로그램 중에서 회계관련 기초프로그램에 참가한 사원과 대리 계층을 중심으로 연구하였다.

2) 교육훈련의 평가수준은 Kirkpatrick(1976)의 평가 4단계 모델에 따라 교육훈련 종료후 측정된 반응, 학습, 전이, 결과를 통틀어 의미한다. 그러나, 이 연구에서는 네 번째 단계인 결과를 제외한 전이까지만 다루었다.

3) 교육훈련의 평가수준 중 세 번째 단계인 전이의 측정은 교육훈련 후 교육훈련 참가자의 상사나 동료의 참가자에 대한 지속적인 관찰이 어렵고, 만약 가능하다고 해도 교육훈련의 결과에 의한 행동변화로 단정하기가 곤란한 점

등이 있어 교육훈련 참가자 자신의 평가로 전이정도를 측정하였기 때문에 다면적인 평가가 되지 못하였다.

회수율을 보였다.

II. 연구방법 및 절차

1. 조사대상의 선정과 표집

가. 조사대상의 선정

교육훈련의 평가수준 중 2, 3단계인 학습과 전이의 측정이 용이할 뿐 아니라 그 결과의 신뢰성과 타당성을 확보하기 위해서는, 가치관이나 태도의 변화를 주목적으로 하는 프로그램(예를 들면, 리더십 개발이나 대인관계 향상 프로그램 등)보다는 직무기술 향상에 목표를 두고 있어 학습의 정도를 측정하기 쉽고 현업에의 적용과 같은 전이 정도도 비교적 분명하게 판단할 수 있는 직무 교육훈련 프로그램으로 조사대상을 삼았다.

따라서 국내의 기업 교육훈련 기관 중 기업 교육훈련에 대한 대표성과 충분한 수의 조사대상자 확보, 그리고 연구자의 추적조사 가능성 등을 고려하여 L 연수원의 회계분야의 5개 기초과정을 조사대상으로 확정하였다.

나. 조사대상자의 선정

5개 프로그램들은 주로 회계분야의 사원들을 대상으로 1996년 3월 4일부터 5월 17일까지 프로그램별로 각각 두 차례에 걸쳐 4박 5일간씩 실시되었다.

총 10회에 걸쳐 실시된 프로그램 중에서 기업 관리회계 실무과정의 2차만을 제외한 모든 프로그램 참가자(845명)를 대상으로 반응, 학습, 전이정도 평가 등 총 3회에 걸친 추적조사를 실시하였다. 그 결과, 조사지에 이름을 기재하지 않았거나 응답이 불성실한 조사지를 제외하고 730명(전체의 86%)의 조사지가 유효하였다. 이 중 전이정도 평가는 우편조사 결과, 총 730명 중 455명의 조사지가 회수되어 62%의

2. 평가수준별 측정방법과 조사도구의 제작

각 평가수준에 대한 조사도구의 제작은 먼저, 관련문헌 연구를 통해 각 변인을 측정하는데 필요한 문항들을 충분히 확보하여 설문으로 구성하고 난 후 석·박사급 3명으로부터 설문에 대한 안면타당도(face validity)를 검토 받았다.

이어서 예비조사(pilot test)를 실시하여 제작된 조사도구의 신뢰도와 타당도를 검증한 후, 그 결과가 좋지 않은 문항들에 대해서는 문항의 내용을 검토하여 수정 혹은 삭제하였다. 신뢰도 검증에는 Cronbach's alpha 계수를 사용하였고, 타당도 검증에는 각 도구의 일개념성(一概念性)을 탐색하기 위해 주축 분해법(principal axis factor method)을 사용하여 공통요인분석(common factor analysis)을 실시하였다.

가. 반응

일반적으로 교육훈련의 반응(reaction)은 교육훈련 프로그램의 모든 측면(예를 들면, 프로그램 설계, 교육훈련 방법, 강사, 교육훈련 환경, 음식 등)에 대한 교육훈련 참가자의 견해나 감정을 측정한다. 그러나 실제 기업교육에서는 교육훈련 프로그램을 마치고 난 직후, 그 프로그램이 실제 업무에 얼마나 도움이 되었는지를 묻는 유용성(usefulness) 파악이 무엇보다 중요한 의미를 가진다. 따라서 이 연구에서도 교육훈련에 대한 반응을 단지 교육훈련 프로그램 전반에 걸친 만족정도가 아닌 교육훈련의 유효성을 평가하도록 조사도구를 구성하였다. Al-ammar(1994)가 최근의 연구에서 '교육훈련 유효성에 대한 교육훈련 참가자들의 인식'(trainees' perceptions of training effectiveness)을 재는 9개의 문항을 개발하였는데 신뢰도 계수도 0.76으로 양호한 편이었다. 이 연구에서도 Al-ammar가 개발한 9개의 문항을 그대로 사용

하기로 하고 예비조사를 실시한 결과, Cronbach's alpha 계수는 0.90으로 높은 편이었다. 본 조사에서도 신뢰도 계수는 0.86으로 비교적 높은 편이었다.

한편, 반응의 일개념성을 측정하기 위해 공

통요인 분석을 실시한 결과, <표 II-1>과 같이 각 문항들이 비교적 서로 유사한 내용을 물어 보고 있었고, <표 II-2>처럼 하나의 개념을 재고 있었다.

<표 II-1> 반응 각 문항에 대한 사전적 공통변량 추정치

문항번호	1	2	3	4	5	6	7	8	9
공통변량	.40	.39	.48	.48	.27	.60	.48	.56	.36

<표 II-2> 반응 문항에 대한 수정상관 매트릭스의 고유치

고유치번호	고유치값	비율	누적비율
1	3.83	.95	.95
2	.62	.15	1.10

나. 학습

교육훈련 참가자의 학습정도는 교육훈련이 끝난 직후 실시된 프로그램별 시험성적으로 측정하였다. 프로그램별 시험문제는 '95년도의 기출문제를 수정·보완하여 사용하였다. 우선, 각 프로그램은 4박 5일간 실시되면서 최소 5-6개 이상의 교과목으로 구성되어 있기 때문에, 이 모든 교과목들의 내용이 시험문제에 고르게 반영되어야 한다. 동시에 프로그램 목표에서 제시하고 있는 각 행동수준(지식, 이해, 적용수준)을 측정할 수 있어야 한다.

따라서 '95년도 시험문제에 대해, 내용*행동 2원 분류표를 작성하여 각 교육훈련 프로그램에서 교육한 내용들이 목표행동 수준(지식-이해-적용단계) 상에서 제대로 평가되도록 문제들이 출제되었는지를 분석하였다. 그 결과, 특정 교과목이나 특정 행동수준에 시험문제가 집중되어 있는 점을 발견하고 삭제할 필요가 있는 문제들을 확인하였다.

다음으로, 기업환경의 변화로 회계관련 기존 법규가 폐지되거나 신규법규가 제정되는 등의 이유로 기출문제 중 시대적 상황에 부합되지 않는 문제를 삭제하거나 신규 문제를 제작할 필요성이 대두되었다.

끝으로 문항분석용 프로그램(ITEMAN)을 사용하여 프로그램별 시험문제에 대한 문항 분석을 실시하여 문항 전체에 대한 신뢰도, 문항 난이도, 그리고 문항 변별도를 파악하였다. 그 결과, 난이도와 변별도가 좋지 않은 문항들을 골라내어 문제 자체를 교체하거나 답지를 수정/교체하는 등의 문제수정을 실시하였다. '95년도 프로그램별 시험문제 수는 총 20문항이었는데 각각 50% 수준인 10문제 정도가 교체 또는 수정되었다. 문항교체나 수정은 프로그램별 내용전문가인 강사들이 직접 수행하였다.

'95년도 시험문제에 대한 문항분석시 사용된 피험자수(표본)는 분석 문항수의 5-10배가 필요하다는 주장(박동건, 1992; p. 41에서 재인용)

에 근접한 3~5배수였다. 아울러 교육훈련 참가자의 부정행위를 방지하기 위해 시험문제의 배열을 바꿔 A형과 B형의 두 가지 형태의 시험지

를 사용하였다. 위와 같은 절차를 걸쳐 작성된 시험문제를 본 조사에 적용한 후 문항분석을 실시한 결과, 아래 <표 II-3>과 같았다.

<표 II-3>

회계분야 5개 기초 프로그램의 문항 분석결과

프로그램	문 제 유 형	표본수	문항수	성 적		척 도 신뢰도	난이도 평 균	변별도 평 균
				평 균	표 준 편 차			
재무회계 1차	A형	50	20	15.1	2.65	0.58	0.75	0.56
	B형	47	20	14.9	3.45	0.77	0.75	0.63
재무회계 2차	A형	84	20	16.1	3.37	0.80	0.81	0.73
	B형	97	20	16.9	2.54	0.66	0.84	0.59
자금관리 1차	A형	41	20	11.9	2.63	0.42	0.60	0.38
	B형	42	20	6.1	1.77	0.24	0.31	0.40
자금관리 2차	A형	36	20	14.4	2.76	0.55	0.72	0.45
	B형	39	20	14.6	2.60	0.52	0.73	0.45
국제금융 1차	A형	23	15	12.7	1.68	0.61	0.84	0.80
	B형	24	15	13.3	1.11	0.21	0.89	0.70
국제금융 2차	A형	26	15	13.1	1.24	0.21	0.87	0.57
	B형	22	15	13.1	1.08	0.07	0.87	0.61
관리회계 1차	A형	67	20	13.8	2.91	0.62	0.69	0.53
	B형	67	20	13.5	2.71	0.56	0.68	0.45
세무회계 1차	A형	41	20	13.8	3.91	0.80	0.69	0.61
	B형	40	20	13.8	3.31	0.69	0.69	0.54
세무회계 2차	A형	50	20	16.6	3.26	0.82	0.83	0.79
	B형	52	20	16.8	2.91	0.77	0.84	0.73

* 국제금융의 경우도 필기시험 15점 만점에 의한 실습점수가 5점 추가되어 총 20점 만점이었음.

<표 II-3>을 보면, 문항분석에 사용된 표본수가 2~3배 수준밖에 되지 못해 충분하지 않았으나, 자금관리와 국제금융 1차의 B형과 국제금융 2차의 A, B형을 제외하고는 척도 신뢰도가 비교적 양호한 편이었다. 또한 난이도 평균은 자금관리 1차의 B형을 제외하고는 모두 0.6에서 0.89에 이르러 이상적인 평균난이도 수준인 0.5(황정규, 1991; 486)를 훨씬 상회하였는데, 이는 각 프로그램이 상대평가가 아닌 절대평가

를 실시하여 12점 이하를 낙제 선으로 간주하고 있기 때문에 문제 출제시 평균난이도를 의도적으로 0.7 수준으로 맞춘 데서 비롯된 결과로 보인다. 그러나 모든 프로그램들의 변별도 평균은 변별도가 있다고 해석하기에 무방한 수준인 0.2 수준(황정규, 1991; 490)을 훨씬 초과하고 있어 전체적으로 우수하다고 판단된다. 이상의 결과를 나타낸 시험성적들을 활용함에 있어 이 연구에서는 다음 세 가지 사항을 검토

하였다.

첫째, 물론 회계분야라는 공통성을 지니고 있기는 하지만, 각 프로그램별 학습목표와 구성 교과목이 상이하기 때문에 동일한 시험성적을 갖고 프로그램간에 동일하게 비교할 수는 없다는 것이다. 따라서 프로그램별로 교육훈련 참가자의 개별성적을 비교해야 된다는 결론을 얻었다.

둘째, 같은 이름의 프로그램 내에서도 차수와 문제 유형별로 평균성적과 문항 난이도 및 변별도 평균이 서로 다르기 때문에 차수와 문제 유형별로 평균성적간에 차이가 있는지를 검증하였다. 그 결과, <표 II-4>와 같이 차수간에

는 차이가 있기도 하고 없기도 했지만, 문제 유형간에는 모두 차이가 없었다. 따라서 같은 프로그램 내에서 차수별로 구분하여 교육훈련 참가자의 개별성적을 상호 비교하기로 하였다.

셋째, 각 프로그램의 차수별로 구분하여 개별성적을 사용할 경우 총 730명 각 개인의 성적을 그대로 서로 비교할 수 없기 때문에, 각 프로그램의 동일 차수 내에서 한 주어진 기준 집단에 대한 한 개인의 위치를 나타내는 백분위(percentile rank)를 사용하였다. 실제로 분석에 투입된 데이터는 백분위를 역연산(100-백분위)한 점수를 사용하였다.

<표 II-4> 프로그램별 차수와 문제유형별 차이 검증

구 분	재 무 회 계	자 금 관 리	국 제 금 융	관 리 회 계	세 무 회 계
차 수 간 F값	14.72**	28.3**	0.56	0.04	31.97**
문제유형간 F값	0.69	0.76	0.15	0.70	0.22

* : $p < .05$ ** : $p < .01$

* 관리회계의 차수간 차이검증 결과는 2차 교육훈련 참가자의 성적을 포함한 것임.

다. 전 이

교육훈련 참가자의 전이는 교육훈련을 통해 학습한 지식과 기술에 대한 '전이수준(level of transfer)'을 측정하였다. 전이수준을 측정하는 설문은 교육훈련을 받고 난 지 3개월이 지난 후에 프로그램에서 배운 지식과 기술을 '현재 직무에 얼마나 많이 사용하고 있는지'를 측정하였다. 동시에 현재 직무에 별로 사용하고 있지 않다면, 그 이유가 무엇인지를 파악하기 위해 그 이유를 7개의 유형으로 나누고 각각에 대한 교육훈련 참가자의 느낌이나 생각 정도를 체크하도록 하였다. 측정방법은 5점 Likert 척도를 사용하였다. 본 연구에서 측정시기를 교육을 받고 난 지 3개월로 정한 이유는 교육훈련의 전이는 교육훈련 종료 후 적어도 3

개월이 지난 후에 측정되어야만 한다는 많은 연구자들의 주장(Al-ammar, 1994;70)에 근거를 둔 것이다.

3. 자료수집

자료수집은 설문지와 시험지를 토대로 예비 조사와 본 조사를 통하여 이루어졌으며, 직접 조사(현장조사)와 우편조사 두 가지 방법을 병행하여 실시하였다. 예비조사는 본 조사에 앞서 설문지의 신뢰도와 타당도, 시험지의 난이도와 변별도 검증을 위해서 실시되었다. 먼저 설문지 검증은 조사대상 연수원에서 '96년 2월 6일부터 10일 사이에 실시된 신입사원 프로그램에 참가한 연수생중 150명을 대상으로 실시

하였고, 시험지 검증은 조사대상 프로그램의 '95년도 시험지 총 648부에 대한 문항분석을 통해 이루어졌다. 본 조사는 예비조사 결과 분석을 토대로 설문지와 시험지를 수정·보완하여 확정된 후, 프로그램별로 총 세차례에 걸쳐 '96년 3월 4일부터 8월말까지 추적조사를 실시하였다. 추적조사는 설문지와 시험지에 소속회사와 이름을 명기하도록 하여 이루어졌다. 이름을 쓰도록 하면 자신의 솔직한 응답보다는 사회적으로 바람직하다고 여겨지는 문항번호에 반응하거나 혹은 불성실하게 응답할 가능성이 크기 때문에 이를 최대한 방지하기 위하여 직접조사시에 조사자가 '이 설문지는 단지 연구와 프로그램 개선에만 활용된다'는 목적을 응답자들에 직접 충분히 설명하도록 하였다.

4. 자료분석 방법

이 연구에서 수집된 자료는 윈도우용 SAS (Statistical Analysis System) 프로그램을 사용하여 분석하였고, 통계적 유의수준은 5%로 하였다. 이 연구에 적용된 통계방법은 기본적으로 빈도와 백분율, 평균과 표준편차를 사용하였고, 교육훈련의 각 평가수준들간의 관계 파악에 상관관계수(Pearson's correlation coefficient)를 사용하였다.

III. 연구결과 및 해석

1. 조사대상자의 일반적 특성

이 연구의 조사대상자로 선정되어 응답한 5개 회계기초 프로그램 참가자들(총 730명)의 일반적 특성을 살펴보면 다음과 같다.

① 성별분포는 남자가 667명(91.4%)으로 절대 다수를 차지하고 있었다.

② 연령과 근속년수는 평균연령이 28.7세, 평균 근속년수가 3.1년으로 회사에 입사한 지 얼마 되지 않은 사람들이 주로 회계관련 기초 프

로그램에 참가했음을 알 수 있었다.

③ 교육수준은 대졸이 580명(79.5%)으로 절대비율을 차지하고 있었고, 직위분포는 사원이 67.7%, 대리가 28.5%로 나타나 두 계층이 전체의 95%를 넘어서고 있었다.

④ 전공은 회계전공이 73.7%로 상대적으로 매우 높게 나타나 기업체의 회계분야 종사자는 학교전공을 고려하여 배치함을 알 수 있었다.

⑤ 교육훈련 참가자들의 현재 담당직무는 기획/관리회계 직무담당자가 35.2%로 가장 많았고, 다음으로 재무회계 담당자(29.6%), 금융(자금)담당자(20.3%) 순이었다.

⑥ 교육훈련 참가자가 이수하게 될 프로그램과 같은 분야에서 근무한 경험이 있는지를 물어 본 결과 '있다'는 응답비율(31.4%)보다 '없다'는 응답비율(68.5%)이 훨씬 많아 프로그램 참가요건에 동일 직무경험을 크게 고려하고 있지 않았다.

⑦ 교육훈련 참가 전에 회계관련 프로그램에 참가한 경험이 있는지에 대해서는 '있다'는 응답비율이 53.6%로 나타나 참가자 중 절반 이상이 회계관련 프로그램에 대한 사전 교육기회가 있었음을 알 수 있었다.

한편, 교육훈련의 각 평가수준은 우선 교육훈련의 유효성에 대한 인식을 측정한 반응평가 점수가 3.45로 이론적 평균치인 3.00보다 높았고, 학습정도를 나타내는 시험성적은 평균 15.23(20점 만점)으로 100점 만점으로 환산할 때 76점 수준으로 비교적 높은 편이었다. 전이수준은 3.09로 낮은 편이었다. 여기서 전이수준에 대한 설문과 연계하여 5점 척도 중 '사용하고 있지 않다'는 응답항목인 1과 2에 체크한 사람들에게 사용하지 못한 이유를 유목화한 문항에 반응토록 한 결과, '나의 현재직무가 내가 이수한 교육과 일치하지 않아서'가 112명의 응답자 중 약 54%인 60명으로 가장 많았고, 그 다음으로는 '내가 배운 모든 것은 현장에 적용할 수 없는 이론적인 것들이어서'가 110명의 응답자 중 약 27%인 30명이었고, 세 번째는 '교육

에서 배운 것을 내 직무에 활용하는데 대한 인센티브가 없어서'가 107명 중 약 21% 수준인 23명이었다. 이러한 결과는 전이수준을 높이기 위해서는 무엇보다 교육훈련의 내용과 현재의 직무가 일치해야 하고, 내용도 이론적인 것이 아니라 현장에 적용 가능한 실제적인 것이어야 하며, 아울러 교육훈련 결과를 제대로 평가하여 인사고과 등에 반영함으로써 실질적인 보상과 연계하지 않으면 안된다는 것을 보여주고 있다.

2. 교육훈련 평가수준들간의 관계 분석결과

연구문제인 '교육훈련 참가자의 반응과 학습간, 학습과 전이간, 반응과 전이간에는 서로 어떤 관계가 있을 것인가?'를 검증하기 위해 설정된 세 가지 가설에 대한 분석결과는 다음과 같다.

가설 1. 교육훈련 참가자의 반응은 학습과 정적인 상관관계가 있을 것이다.

교육훈련 참가자의 반응과 학습은 상관계수가 0.11($p < .01$)로 유의한 상관관계가 있었기 때문에 이 연구가설은 채택되었다. 그러나 상관계수를 활용하여 반응의 학습에 대한 설명 및 예측변량(R^2)을 구한 결과, 1.2% 정도로 극히 미약하였다.

가설 2. 교육훈련 참가자의 학습은 전이와 정적인 상관관계가 있을 것이다.

교육훈련 참가자의 학습과 전이수준은 상관계수가 0.10($p < .05$)으로 유의한 상관관계가 있어 이에 대한 연구가설은 채택되었으나, 학습의 전이수준에 대한 설명 및 예측변량(R^2)은 1.0% 정도로 매우 미약하였다.

가설 3. 교육훈련 참가자의 반응은 전이와 정

적인 상관관계가 있을 것이다.

교육훈련 참가자의 반응과 전이수준은 상관계수가 0.23($p < .01$)으로 유의한 상관이 있어 이 연구가설은 채택되었다. 상관계수를 활용하여 반응의 전이수준에 대한 설명 및 예측변량(R^2)을 구한 결과, 5.3%였다.

이상의 분석 결과, 교육훈련 참가자의 반응과 학습간, 학습과 전이간의 상관계수가 통계적으로는 유의하였지만, 그 정도가 매우 약했다. 아울러, 독립변인의 종속변인에 대한 설명 및 예측변량 또한 매우 미미한 것으로 나타났다. 그러나 반응과 전이간의 상관계수는 상대적으로 높게 나타났는데, 이는 전이수준을 반응평가와 같은 자신의 인식정도를 묻는 자기평가 방식으로 측정된 결과일 수도 있다.

IV. 결론 및 제언

우리 기업의 교육훈련 수준이 과거와 비교해 볼 때 양과 질적인 측면에서 놀랄만한 발전을 이루긴 했지만, 교육훈련 평가는 그 대부분이 교육훈련 참가자의 프로그램에 대한 만족 정도를 측정하고 있고, 학습 정도는 일부에서 실시하고 있을 뿐, 교육훈련을 받고 난 후 그 결과가 실제 현업에서 얼마나 활용되고 있으며 업무 및 조직의 생산성에 어떤 영향을 미치고 있는지에 대해서는 거의 평가가 이루어지고 있지 않다.

이런 상황에서 교육훈련의 반응과 학습, 학습과 전이간에 약한 상관이 존재하고, 위계적 관계를 나타내는 설명 및 예측변량은 극히 미약하게 나타난 이 연구의 결과를 보면, Kirkpatrick(1976)의 반응→학습→전이에 대한 이론적인 위계적 모델이 강하게 지지받을 수 없을 뿐 아니라, 선행 평가수준이 필요조건임을 충족시키지 않기 때문에 즉, 만족스런 반응으로 충실한 학습을 가정하기 어렵기 때문에 각 평가수준별로 별도의 측정이 이루어져야 함

을 시사한다.

이와 같은 결과를 다른 측면에서 보면, 평가 수준별로 관련있는 혹은 원인이 되는 요인들이 서로 다르기 때문인 것으로도 볼 수 있다. 다시 말하면, 반응의 경우는 교육훈련 참가자 개인의 심리적 특성에 크게 좌우되는 반면, 학습의 경우는 교육훈련 참가자의 개인적 배경 예를 들면, 전공, 관련분야에 대한 사전적 교육 기회 및 직무경험 등이 심리적 특성보다 더 크게 영향을 미치기 때문에 위계적 연결고리가 약할 수도 있다는 것이다.

앞으로 보다 나은 연구를 위해 이 연구를 수행하면서 느꼈던 제한점을 중심으로 몇가지 사항을 제언하고자 한다.

1) 이 연구에서는 교육훈련의 평가수준을 3 단계인 전이까지만 연구대상으로 삼았는데 향후에는 유효성의 4단계인 조직의 생산성, 이직율, 직무만족 등과 같은 결과(result) 수준의 평가도 포함하여 교육훈련 평가수준 전체의 위계성을 실증적으로 분석하여 검증할 필요가 있다.

2) 교육훈련의 평가수준을 측정하는 방법에는 여러 가지가 있을 수 있는데, 이러한 방법들간에 상호 어떤 관계성을 갖고 있는지를 체계적으로 분석하고 이를 교육훈련 현장에 실용화할 수 있도록 노력해야 할 것이다. 다시 말하면, 반응평가의 경우 교육훈련의 유효성에 대한 인식으로 측정하지 않고, 교육훈련 전반에 대한 행복감이나 프로그램 설계에 대한 만족도로 측정했을 때에는 그 결과는 서로 다를 수 있다(Goldstein, 1986). 따라서, 반응을 측정하는 여러 가지 형태의 측정방법을 상호 비교하여 그들간에 어떤 연관성을 갖고 있는지를 분석할 필요가 있다. 학습평가의 경우도, 시험문제 등을 통해 객관적으로 평가하는 경우와, 참가자가 선수학습 정도가 상이하여 출발점 행동이 서로 다른 조건이므로 시험문제 등을 통해 교

육훈련 전의 학습정도와 교육훈련 후의 학습정도 간의 차이를 재는 경우, 혹은 설문서 등을 통해 많은 것을 배웠는지에 대한 개인적 인식으로 학습정도를 측정하는 경우 교육훈련 유효성 연구의 결과가 서로 같은지, 혹은 다른지를 비교 분석할 필요가 있다. 또한, 지식이나 기술에 초점을 둔 프로그램 평가와 태도나, 가치, 관점을 변화시키는 프로그램 평가는 내용과 방법에 있어 서로 다르기 때문에 둘 간의 교육훈련 유효성 연구결과도 어떠한지를 비교 분석할 필요가 있다. 전이수준 평가도 마찬가지로, 개인의 인식에 기초한 평가(self-report)와 직속상사나 동료에 의한 평가 간에도 어떤 연관성을 갖고 있고, 교육훈련 유효성의 연구결과에는 어떻게 작용할 지에 대한 비교 분석이 필요한 것이다.

V. 참고 문헌

1. 권대봉. 1993. 납득할 수 있는 교육효과의 측정 방법, 인사관리 통권 43호 '93(3), p. 10.
2. 박동건. 1992. '심리측정의 개관: 신뢰도와 타당도를 중심으로', 심리검사 제작의 이론과 실제, 한국심리학회.
3. 윤영득. 1994. 한국 기업교육의 대안적 평가모델에 관한 연구: ROI 평가를 중심으로, 한양대학교 석사학위 논문.
4. 이도형. 1995. 조직내 교육훈련의 학습 및 전이 효과 모델 검증연구, 산업 및 조직심리학회 동계 학술발표대회 논문집, pp. 17-54.
5. 이순목. 1992. '문항분석과 검사타당화', 심리검사 제작의 이론과 실제, 한국심리학회.
6. 이종원, 최현집. 1996. SAS를 이용한 통계분석, 박영사.
7. 황정규. 1991. 학교학습과 교육평가, 교육과학사.
8. Al-ammam, S. A. 1994. The Influence of Individual and Organizational Characteristics on Training Motivation and Effectiveness, Unpublished Dissertation, the State Univ. of New York at Albany.
9. Alliger, J. M. and E. A. Janak. 1989. Kirkpatrick's

- Levels of Training Criteria: Thirty Years Later, *Personnel Psychology*, 42:331-341.
10. Clement, R. W. 1978. An Empirical Test of the Hierachy Theory of Training Evaluation, Unpublished Doctoral Thesis, Michigan State University.
 11. Faerman, S. & C. Ban. 1993. Trainee Satisfaction and Training Impact: Issues in Training Evaluation, *Public Productivity Management Review*, 16:299-314.
 12. Fleishman, E. A. 1955. Leadership Climate, Human Relations Training, and Supervisory Behavior, *Personnel Psychology*, 6, 205-222.
 13. Fromkin, H. L., J. Brandt, D. C. King, J. J. Sherwood & J. Fisher. 1975. An Evaluation of Human Relations Training for Police, *Catalog of Selected Documents in Psychology*, 5, 206-207.
 14. Goldstein, I. L. 1986. Training in Organization: Needs Assessment, Development and Evaluation, Monterey, CA: Brooks/Cole.
 15. Hamblin, A. 1974. *Evaluation and Control of Training*, London: McGraw-Hill.
 16. Kirkpatrick, D. L. 1959. Techniques for Evaluating Training Programs, *Training and Development Journal*, 13:3-9.
 17. Kirkpatrick, D. L. 1976. Evaluation of Training, in R. L. Craig(ed.), *Training and Development Handbook*(2d ed.), N.Y.: McGraw-Hill.
 18. Kirkpatrick, D. L. 1987. Evaluation, in R. L. Craig (ed.), *Training and Development Handbook*(3rd ed.), New York: McGraw-Hill Book Co..
 19. Kirkpatrick, D. L. 1994. *Evaluating Training Programs: The Four Levels*, Berrett-Koehler Publishers, Inc..
 20. Latham, G. P., K. N. Wexley & E. D. Purcell. 1975. Training Managers to Minimize Rating Errors in the Observation of Behavior, *Journal of Applied Psychology*, 60, 550-555.
 21. Mathieu, J. E., S. I. Tannenbaum & E. Salas. 1992. Influence of Individual and Situational Characteristics on Measures of Training Effectiveness, *Academy of Management Journal*, Vol. 35, No.4:828-847.
 22. Noe, R. A., & N. Schmitt, 1986. The Influence of Trainee Attitudes on Training Effectiveness, *Personnel Psychology*, 39:497-523.
 23. Trefz, M. K. 1991. Factors Associated with Perceived Efforts of Trainees to Transfer Learning from a Management Training Activity, Unpublished Dissertation, Ohio State University.