



4차 산업혁명시대의 교육 효과 및 성과 측정방법

두하우컨설팅 정하늘 컨설턴트

당신은 준비된 HRDer입니까?

코로나 19 사태가 장기화됨에 따라 기업을 둘러싼 대내외 환경이 빠른 속도로 변화하고 있다. 많은 기업들이 재택근무와 시차출퇴근제, 원격근무 체제를 전격 도입하는 한편, 회의실과 탕비실, 헬스시설 등 각종 공용공간의 사용을 폐지하였으며, 감염 확률을 낮추기 위해 제각기 칸막이와 파티션을 설치했다. 학자들은 포스트 코로나 시대를 위한 대책들을 내놓았고 비대면 서비스를 제공하는 업체들은 환호성을 질렀으며 유통업체들은 과로에 시달렸다. 누군가에게는 희망이, 또 누군가에게는 절망이 가득한 한 해였다.

아마 이 깊은 수렁과도 같았던 한 해를 어렵사리 헤쳐나온 사람들 중 하나가 바로 HRD 종사자들이 아닐까 싶다. 2020년 6월, IBK경제연구소의 조사 결과에 따르면, 중소기업 응답자 중 82.0%가 코로나 19 이후 경영상황이 나빠졌다고 응답했으며, 특히 교육 서비스업·기타 개인 서비스업 종사자의 응답률이 100.0%에 달하는 것으로 나타났다. 이러닝의 비율이 늘고 있다고는 하나, 아직까지 집합교육의 비중이 상당했던 코로나 이전의 교육 국면을 생각하면 당연한 결과라 할 수 있을 것이다.

<표 1> 코로나19가 중소기업에 미친 영향

산업 대분류	피해를 입음(%)	피해를 입지 않음(%)
전체 산업	82.0	18.0
제조업	74.3	25.7
교육 서비스업	100.0	0.0
협회 및 단체, 수리 및 기타 개인서비스업	100.0	0.0
숙박 및 음식점업	98.5	1.5
부동산업	94.0	6.0
도매 및 소매업	68.1	31.9
전문, 과학 및 기술 서비스업	56.2	43.8

• 출처 : 시장경제. (2020.06). 中企 82% 코로나 피해... 교육서비스업, 가장 타격 컸다.



그럼 코로나가 비단 HRD 시장에 좋지 않은 영향만 끼쳤는가? 단기적으로보면 우리 HRDer들에게 여러 과제들을 안겨준 건 사실이다. 하지만 조금만 깊게 생각해보면 언젠가 반드시 해결해야할 문제였다는 걸 깨닫는 건 그리 어렵지 않다. 코로나는 단지 그 시기를 획기적으로 앞당기는 트리거가 되었을 뿐, 우리가 지난 해 마주한 역경과 당황은 어쩌면 우리가 미래 교육의 형태, 미래 교육의 전달방식에 그만큼 준비되어있지 않았다는 반증이기도 하다.

그럼 앞으로 우리는 무엇을 준비해야하는가? 어떠한 철학과 어떠한 무기를 가지고 우리의 소임을 다하며 역할을 증명할 것인가? 그 수단 중 가장 강력한 하나가 될 수 있는 교육 효과 및 성과 측정 방법의 과거와 현재, 그리고 미래에 대해 논해보고자 한다.

어떻게 평가해왔는가?

지금까지 우리는 교육을 어떻게 평가해왔는가? 교육 평가와 관련하여 다양한 이론과 모델들이 존재하지만 대체로 결과를 중심으로 한 접근 방식과 과정을 중심으로 한 접근 방식으로 분류가 가능하다¹⁾. 결과를 중심으로 한 평가 방법에는 Kirkpatrick 4단계 모형과 Phillips의 5단계 모형 등이 포함되며, 과정 중심 평가 방식에는 CIPP, CIRO 모형 등이 포함된다. Kirkpatrick 4단계 모형과 같은 결과 중심 평가는 교육을 통해 어느 정도의 목표가 실현되었는지 규명하는데 중점을 두며, CIPP와 같은 과정 중심 평가는 학습 활동의 가치를 높이는데 영향을 미치는 과정적 요인들과 그 요인간의 인과 관계를 밝히는 데 초점을 둔다.

<표 2> 전통적인 교육 평가 방법론

구분	Kirkpatrick 4단계 모형 ²⁾	CIPP ³⁾
접근 방식	결과 중심	과정 중심
구성요소	 <ul style="list-style-type: none"> • 1수준 : 반응평가(Reaction) • 2수준 : 학습평가(Learning) • 3수준 : 행동평가(Behavior) • 4수준 : 결과평가(Result) 	 <ul style="list-style-type: none"> • 요소1 : 상황평가(Context) • 요소2 : 투입평가(Input) • 요소3 : 과정평가(Process) • 요소4 : 산출평가(Product)

<표2>의 대표적인 교육 평가 모델들을 보면 더 넣고 말것도 없이 훌륭한 모델들로 보인다. 그럼 대체

1) Kraiger et al. (1993).

2) Kirkpatrick. (1959).

3) Stufflebeam & Shinkfield. (2007).

왜 이 훌륭한 모델들을 활용하지 않는가? 교육 평가 업무를 한 번이라도 경험해본 사람이라면 알겠지만 사실상 저 모든 요소들을 측정하기란 쉬운 일이 아니다. 아마 대다수의 조직에서 사용하고 있을 Kirkpatrick 모형을 예로 들어보자. 1단계 반응 평가는 교육 종료 직후 서베이 한 번이면 끝이 난다. 가장 쉽고 간단해 현업자들도 가장 많이 활용하는 평가 방식 중 하나이다. 하지만 문제는 2단계부터 시작된다. 참여자들의 지식(Knowledge), 기술(Skill), 태도(Attitude)의 변화를 어떻게 측정할 것인가? 많은 HRDer들이 이에 '지필평가'를 택한다. 하지만 지필평가는 논의의 여지가 많다. 일단 지필평가로 학습자의 변화를 측정할 수 있는 교육이 제한적이거나, 두 번째로 지필 평가는 학습자의 지식 변화는 측정할 수 있지만 기술과 태도 변화를 측정하는데 적합한 도구는 아니기 때문이다.

3단계 평가도 만만치는 않다. 통상적으로 학자들이 적절하다고 주장하는 행동 평가의 시기는 교육이 종료된 후 3~6개월 정도가 지난 시점이다. 일단 3~6개월 전에 실시한 교육 참여자를 추적하여 평가를 실시해야 한다는 점부터가 고난의 시작이다. 교육 직후 측정한다 하더라도 학습자의 행동을 일일이 관찰하는 방법을 쓰지 않는 이상 결국 행동 변화 여부는 학습자의 자기보고식(self-reporting) 응답을 수집할 수 밖에 없는데 이러한 응답 방식은 객관성의 문제가 있다. 그래도 3단계까지는 어떻게든 정량적 측정이 가능하다. 4단계, 교육이 조직의 성과와 목표에 어떤 영향을 미쳤는가? 이 질문은 정량적 근거 제시가 어렵다. 정성적으로 달성 정도를 제시할 수 있지만 교육 예산을 수립하는, 숫자를 좋아하고, 보다 객관적인 증거를 원하는 간간한 경영진들에겐 그다지 효력을 발휘하지 못한다.

어떻게 평가하고 있는가?

그렇다면 실제로 국내 우수 대기업들은 어떻게 학습 평가를 하고 있을까? 그들에게는 특별한 평가 도구가 존재하는가? 해답을 찾기 위해 인터뷰를 실시했다.

<표 3> 국내 민간기업 교육평가 현황 조사 개요

구분	내용
조사 개요	국내 대기업의 교육평가 현황을 조사하기 위해서 인터뷰 질문지를 개발하고 이를 활용해서 전화 인터뷰를 시행함.
조사 대상	연구소, 통신회사, 중공업, 화학, 그룹연수원의 HRD 담당자
조사 기간	2020.05.11(월) ~ 05.15(금)
인터뷰 문항	<ul style="list-style-type: none"> 교육과정에 대한 평가는 어떻게 진행하고 있습니까? (활용모델 등) 커크패트릭 평가모델을 활용하고 있습니까? 활용하고 있다면, 커크패트릭 평가모델의 수준별 평가비율은 어떠합니까? (전체과정 대비 평가비율) ROI 또는 경영성과 기여도를 평가한 경험이 있습니까? 있다면, 어떤 프로세스로 평가를 진행하였습니까? 그리고 평가 결과는 어떻게 분석되었습니까?

조사 결과, 인터뷰 기업 모두 교육 평가에 Kirkpatrick의 4단계 모형을 활용하고 있으며, 활용수준의 차이는 있지만 주로 반응평가, 학습평가 위주로 활용하고 있는 것으로 나타났다. 행동평가는 활용

비율이 낮았고, 결과평가는 활용하고 있지 않았다. 결과평가를 활용하고 있지 않은 이유에 대해서는 결과평가를 위해 소요되는 시간, 자원 등의 투입 노력과 비교하여 결과에 대한 인정이나 신뢰도가 확보되지 않기 때문이라고 응답하였다. 또한, 교육부서의 역할이 현장에서 활용될 수 있는 문제해결이나 비즈니스 모델 개발, informal learning 확대 쪽으로 변화하고 있어 과정 특성상 평가의 중요성이 높지 않게 여겨진다는 의견도 있었다.

<표 4> 조사기업별 평가현황 정리

구분	반응평가 (1단계)	학습평가 (2단계)	행동평가 (3단계)	결과평가 (4단계)
연구소	100%	70%	10%	-
통신회사	10%	-	-	-
중공업	100%	5%	-	-
그룹연수원	100%	50%	10%	-
화학	100%	50~60%	20%	-

<표 5> 조사 기업 Real Voice

Q : Kirkpatrick 모형을 사용하시는 이유가 있을까요?

- 기업 1 : 아무래도 대표적인 평가 모형이기 때문이겠지요? (일부 단계이기는 하지만) 실제 진행하기에 가장 용이하기도 하구요. 일단 다른 모형에 비해서 명확하게 교육과정의 성과를 제시할 수 있어서 자주 사용하는 것 같습니다.
- 기업 2 : 반응평가가 학습 종료되는 시점에 가장 다양하게 교육에 대한 정보를 얻을 수 있으니까요. 이것 활용하면 결과보고하기도 좋고 교육과정 개선에도 직접적으로 활용할 수 있죠.

Q : 결과 평가는 왜 하지 않으시나요?

- 기업 1 : 들이는 시간에 비해 평가결과에 대한 인정을 받기도 어렵고 아무래도 완전히 신뢰롭다고 보기도 어렵구요.
- 기업 2 : 경영성과 기여도와 ROI를 측정하려면 1~3단계 평가 모두 시행하고 현장 자료를 취합해서 다양한 추가 활동(자료조사, 인터뷰, 설문 등)을 수행해야하는데 아무래도 지원을 얻기가 쉽지는 않죠.
- 기업 3 : 기업 내 교육부서의 역할이 현장에서 활용될 수 있는 문제해결이나 비즈니스 모델 개발, informal learning 쪽으로 넘어가고 있어서 실질적으로 평가를 복잡하게 해야하는 필요성을 느끼지 못해요.

왜 평가가 필요한가?

그럼 평가는 필요치 않는가? 필자는 필요하다고 생각한다. 코로나 팬데믹이 도래한 이후 우리는 어떤 식으로 교육 전달 수단을 변화시키고 내용을 최적화시킬지 고민해왔을 뿐, 그에 따라 변화되어야 하는 교육의 체제적 측면에 대해서는 심도 있는 논의가 이루어지지 않아왔다. 특히 경영진 측면에서 '그럼에도 불구하고' 교육에 투자해야 하는 이유를 설명해줄 수 있는 효과 측면은 더욱 그러했다.

요즘같이 변화의 속도가 빠르고 HRD에 있어 애자일 전략이 중시되는 환경에서 평가는 더 이상 필요하지 않을 수 있다. 경영진과 사회가 요구하는 대로 빠르게 개발해서 빠르게 운영하면 되니 소모적 측면이 크고 장기적 측정이 어려운 평가는 중요하지 않다는 논지이다. 정말 그럴까? 그 답을 알아내기 위해 우리는 교육 평가의 정의와 목적을 다시 돌이켜 볼 필요가 있다.

교육 평가. 가장 대중적으로 사용되는 단어이기에 학문적으로 따지자면 분야와 학자에 따라 무수한 정의들이 있겠지만, 일반적으로 인재·조직개발 분야에서의 평가는 특정 제도, 활동, 프로그램 등의 시행이 의도한 목표를 달성했는지 판단하는 활동을 의미한다⁴⁾.

이야기를 진행하기 전에 한 가지 짚고 넘어갈 것은 필자가 여기서 논하고자 하는 교육 평가의 범위가 HRD의 모든 실천 영역을 아우르는 평가는 아니라는 점이다. HRD는 그 오랜 학문의 역사 만큼이나 조직의 제도적 측면까지 넓은 범위의 실천 형태를 포괄할 수 있기 때문에 전체적인 직무 성과나 제도, 기타 지원 활동과 연관된 모든 평가 방식을 이 짧은 글에서 논하기에는 어려움이 있다. 필자가 여기서 논하고자 하는 교육 평가의 범위는 철저히 교육 프로그램 중심이다. 4차 산업혁명의 도래와 코로나 팬데믹에 의해 가장 역동적인 변화가 발생한 실천 영역이 바로 교육 프로그램 영역이기 때문이다.

교육 프로그램 분야의 평가 목적은 크게 2가지로 나뉜다. 1) 해당 프로그램에 의해 실제 학습자의 지식, 기술, 태도가 의도하는 수준으로 상향되었는가? 2) 해당 프로그램은 그 자체로서의 질을 담보하고 있는가? 1)의 질문의 경우 프로그램의 실질적인 효과성에 대한 답을 얻을 수 있고, 2)의 질문의 경우 효과가 있다면, 구체적으로 어떤 부분이 개선되어야 하는가에 대한 답을 얻을 수 있다. 그러나 평가를 수행하지 않을 경우 우리는 이 프로그램이 실제로 의미가 있는지, 어떤 방식으로 개선해야 나가야하는지에 대한 방향을 알 수 없다. 그저 필요하다 외치는 것을 교육으로 기획하고 그것이 괜찮았는지 아닌지에 대한 검증도 하지 못한 채 다시 새롭게 필요하다 외치는 것을 위한 교육을 만들기를 반복할 것이다. 평가는 필요하다. 단지, 그 속도를 높여야 할 필요성에 직면했을 뿐이다.

어떻게 평가할 것인가?

그렇다면 앞으로의 평가는 어떻게 변화해야 할 것인가? 몇 년 전, AI를 이용해 교육 어플을 개발해 화제가 된 기업이 있다. '산타토익'이라는 기업이다. '산타토익'은 학습자의 오답

4) 김태성, 장지현, 백평구. (2019). 조직의 가치와 목표 중심의 HRD 평가. 박영Story.

성향을 시가 분석해 유사한 유형의 문항을 학습자에게 지속적으로 노출시킴으로써 학습의 효과성을 제고시켰다. 누군가는 이것을 큐레이팅의 위력이라 말할 것이고, 누군가는 이것을 인공지능의 승리라 말할 것이다. 필자는 이를 시간과 자원의 한계로 인해 제한적 도입만 이루어져 왔던 완전학습(Mastery Learning)⁵⁾ 개념의 기술적 부활이라 평하고 싶다. 예전에는 불가능했던 이론들이 AI, 빅 데이터 분석 기술의 발전으로 인해 구현 가능한 시대가 도래한 것이다.

필자는 평가 또한 위와 유사한 방식으로 변화해야 한다고 생각한다. 평가에 들어가는 자원 대비 효율성과, 평가의 속도가 문제라면 결국 그걸 해결할 수 있는 가장 좋은 대안은 축적되어 있는 '데이터'를 HRD관점에서 재해석하고 그 처리를 '자동화'시키는 것이다. 예를 들어 당신이 현재 담당하고 있는 교육 프로그램의 학습 몰입도를 측정해야 한다고 생각해보자. 코로나 이후 진행되고 있는 교육들은 대부분 비대면 방식이니 실시간으로 Zoom을 활용했다고 가정해보자. 과연 어떤 데이터들을 평가해 활용할 수 있을까?

필자라면 교수자와 학습자에게 Zoom의 손들기 기능을 활용할 것을 권장한 뒤, 이 손들기 기능이 해당 강의에서 어느 정도의 빈도로 기능하였는지를 Reporting 하여 평가 지표로 활용할 것이다. 초반에는 이 기능이 실제 성과와 어느 정도의 연관성이 있는지 자기보고식 결과와 어떤 관계를 가지는지 여러가지 검증이 필요하겠지만, 만약 이 값을 안정적으로 활용할 수 있다고 한다면 더 이상 반응평가를 위해 설문을 돌리는 수고로움을 겪지 않을 수 있을지도 모른다. 중요한 것은, 지속적으로 시도하는 것이다. 어떤 데이터를 가져올 수 있는가? 어떤 데이터를 활용할 수 있는가? 이 데이터와 기존 요인과의 관계는 어떠한가? 이 데이터는 기존 측정 항목을 대체할 수 있는가? 끊임없는 질문만이 굳어버린 사고를 확장시키고 우리를 새로운 세계로 데려다줄 수 있다. Kirkpatrick 이후 멈춰버리다시피한 교육평가에 새로운 바람이 불기를 기원하며 본 고를 마친다.



- 저자 : 정하늘 컨설턴트
- 두하우컨설팅 블로그 : <https://blog.naver.com/dohow123>
- 두하우컨설팅 홈페이지 : <http://www.dohow.co.kr/>

5) 미국 교육심리학자 벤자민 블룸(Benjamin S, Bloom)에 의해 주창된 학생 지도 전략. 학생들이 차시 학습으로 넘어가기 전 그 수준의 정보에 통달하는 단계(지식테스트의 90%에 통달하는 단계)에 이를 때까지 반복하여 지도해야 한다는 관점으로 학습에 요구되는 시간이 많고, 모든 학생들이 해당 수준의 학습을 성취할 수 있는 조건을 갖추기 위해 교수자들이 학생들에게 개인화된 학습환경을 제공해야했기 때문에 교수자의 많은 사전 연구가 필요하다.