

Rasch 모형을 활용한 관리자 코칭 행동 측정도구 재검증 연구*

이윤수**, 조대연***

■ 요약 ■

본 연구는 국내 관리자 코칭 행동과 관련한 실증연구에서 가장 많이 사용된 조대연, 박용호(2011)의 검사도구를 문항반응이론에 기초한 Rasch 모형을 활용하여 문항 적합도, 곤란도, 응답 범주의 적절성 측면에서 타당한지 다차원적으로 재검토하고, 보다 객관적으로 검사도구의 개선점을 제공하고자 하는 목적을 가지고 있다. 이에 본 연구는 기존의 요인분석 방법뿐만 아니라 Rasch 모형을 다점 척도로 확장한 평점척도모형을 활용하여 관리자 코칭 행동 검사도구를 다각도로 재검증하였다. 분석 결과, 첫째, 현재 관리자 코칭 행동 검사도구는 문항 곤란도가 수정될 필요가 있고, 둘째, 2요인 구조가 적합한지에 대한 이론적·실천적 추가 논의가 요청되며, 셋째, 문항 적합도가 부적절해 보이는 문항에 대한 개선이 요구되고, 넷째, 문항 응답 범주에 대한 합리적인 의사결정이 필요한 것으로 분석되었다. 분석 결과를 바탕으로 본 연구는 인적자원개발 분야에서 이론적·실천적 시사점에 대해 논의했고, 후속연구에 대해 제언 및 Rasch 모형을 활용한 분석 검증 노력의 필요성에 대해 강조하였다.

[주제어] : 관리자 코칭 행동, Rasch 모형, 평점척도모형

* 본 논문은 고려대학교에서 지원된 연구비로 수행되었음.

** 제1저자, 고려대학교 교육학과 연구교수. leoyunsoolee@gmail.com

*** 교신저자, 고려대학교 교육학과 교수. chodae@korea.ac.kr

I. 서론

1. 연구의 필요성

최근 코칭에 대한 국내 연구는 매우 활발하게 이루어지고 있다. 조성진, 정이수(2018)의 코칭 동향 연구에 따르면, 국내 코칭 관련 연구는 2017년까지 총 326편이 출판되었고, 코칭, 교육학, 경영학, 심리학/상담, 복지학/아동 등 다양한 분야에서 연구되고 있으며, 대부분(80% 이상)이 2010년 이후로 출판되는 등 지난 10년간 많은 연구가 이루어지고 있음을 알 수 있다. 성세실리아, 조대연(2016)은 인적자원개발 분야에 초점을 두고 코칭 관련 국내 학술지 연구의 동향을 탐색했는데, 2015년까지 총 53편의 연구물이 출판되었고, 주로 코칭의 효과에 대해 분석했음을 발견했다. 관리자 코칭 행동을 포함한 코칭이 인적자원개발 분야를 포함한 유사 분야에서 꾸준히 주목받는 이유는 성세실리아, 조대연(2016)의 연구에서 유추할 수 있듯이 코칭의 효과성이 입증되었거나 입증하려는 노력에 힘입어 꾸준히 이어져오고 있기 때문으로 보인다. 특히 관리자의 코칭 행동은 조직원들의 직무만족과 성과뿐만 아니라(Ellinger, Ellinger, & Keller, 2003) 경력성공에도 기여하는 것으로 알려져 있다(이운수, 조대연, 김재민, 2018).

관리자 코칭은 직원의 성과와 직무기술을 향상시킬 수 있는 기회를 제공하는 일상적인 업무 과정으로서 리더십의 독특한 스타일로 간주된다(Ellinger, Ellinger, & Keller, 2005). 관리자 코칭 행동은 관리자가 코치로서 제공하는 구체적인 지원 행동을 의미하는데(이운수, 조대연, 김재민, 2018), 일부 학자들은 코칭 리더십이라는 용어를 사용하기도 한다. 관리자 코칭 행동은 Ellinger 외 (2005)의 정의에서 리더십의 한 스타일로 정의된다는 점, 관리자 코칭 행동과 코칭 리더십의 발현 주체가 리더 혹은 관리자라는 점, 코칭 리더십의 정의와 문항이 행동으로 제시되어 있다는 점에서 본 연구에서 관리자 코칭 행동과 코칭 리더십은 동일한 개념으로 간주하고, 관리자 코칭 행동이라는 용어로 통일하여 사용한다. 조대연, 박용호(2011)는 관리자 코칭 행동으로 문제해결력 향상을 위한 지원(이후 문제해결력 향상)과 학습촉진을 통한 역량향상지원(이후 학습촉진)을 제안했고, 강영순, 김정훈(2013)은 방향 제시, 역량 개발, 수행 평가, 관계의 4가지 요소를 제안했으며, 조은현(2011)은 믿음, 존중, 목표설정과 피드백, 질문의 4가지 요소를 제안했다. 학자마다 관리자 코칭 행동을 구성하는 요인은 다르게 보았지만 관리자 코칭 행동의 정의와 목적은 리더 또는 관리자가 구성원의 성장과 학습을 촉진하기 위한 행동이라는 공통점이 있다.

오효성(2009), 조대연, 박용호(2011), 조은현(2011)은 관리자 코칭 행동을 측정할 수 있는 검사도구를 혼합연구(인터뷰 또는 델파이 조사 + 설문조사)를 통해 직접 개발했다. 오효성(2009)과 조은현(2011)의 검사도구 개발과정은 문헌분석, 코칭 전문가 또는 코칭을 받은 경험이 있는

직장인들을 대상으로 한 심층인터뷰, 초점집단인터뷰, 회사원들을 대상으로 한 설문조사 등의 절차를 통해 개발되었고, 요인분석 등을 통해 검증되었다. 조대연과, 박용호(2011)의 연구의 선행연구인 조대연, 김희영(2009)의 연구는 코칭 프로세스에서 코칭 행동을 도출했다는 점에서 코칭 역량 스킬에 대한 문헌분석에서 예비문항을 도출한 오효성(2009)의 연구나 개방형 질문을 통해 예비문항을 도출한 조은현(2011)의 연구와 차별화된다. 강영순, 김정훈(2013)은 외국 학자(Stowell)의 검사도구를 번안해서 사용했는데, 측정도구 개발 및 타당화가 연구의 주요 목적은 아니었다. 2019년 2월 학술연구정보 서비스(RISS)에서 관리자 코칭 행동을 키워드로 검색했을 때 나타나는 국내학술지논문은 총 31편, 학위논문은 66편인데, 문헌분석, 질적연구, 중복 검색된 논문 등을 제외한 실증연구에서는 모두 상술한 네 개의 검사도구가 사용되었고, 조대연과 박용호(2011)의 측정도구를 사용한 연구는 총 23편으로 가장 빈번하게 사용되었음이 확인되었다.

그런데 집계된 23편의 연구물 중 이윤수, 조대연, 김재민(2018)의 연구를 제외하고는 요인분석이 제대로 수행되지 않았거나 측정문항이 온전히 사용된 경우는 없었다. 박근수와 김주후(2011)의 지적처럼 코칭과 관련한 측정도구들은 외국 학자들이 개발한 척도를 번안하고 수정하는 수준에서 벗어나지 못했거나 조성진, 정이수(2018)의 지적처럼 객관적으로 검증된 측정도구를 사용하지 않은 연구를 쉽게 찾아볼 수 있다. 한편, 검사도구 개발 과정을 살펴보면 상관분석에 기초한 요인분석에 의존했다는 점에서 Rasch 모형과 같이 문항반응이론에 기초하여 평가하고 검증하려는 시도가 부족했었다. 문항반응이론은 문항의 특성과 연구 대상의 능력을 독립적으로 추정하고 문항의 신뢰도와 연구 대상의 신뢰도를 추정하는 장점이 있다(박은영, 박소연, 2008; Hong & Wong, 2005). 이에 본 연구는 조대연, 박용호(2011)의 관리자 코칭 행동 측정도구의 후속연구로서 기존 요인분석 방법뿐만 아니라 문항반응이론에 기초한 Rasch 모형을 활용하여 문항 적합도, 곤란도, 응답 범주의 적절성 측면에서 타당한지 다차원적으로 재검토하고, 보다 객관적으로 검사도구를 개선하기 위한 시사점을 제공하고자 한다.

II. 이론적 배경

1. 관리자 코칭 행동 검사도구

조대연, 박용호(2011)의 측정도구는 조대연, 김희영(2009)의 연구에서 출발했다. 조대연, 김희영(2009)은 기업에서 성공적인 코칭을 위해 전문코치들이 수행해야 할 행동들을 분석했는데, 문헌분석을 통해 4가지 코칭 프로세스와 각 프로세스별 36가지 코칭 행동(계약 단계 코칭 행

동 4가지, 자료수집·분석 및 피드백 단계 코칭 행동 12가지, 코칭 실행 및 추수지도 단계 코칭 행동 16가지, 평가 단계 코칭 행동 4가지)을 도출했다. 다음으로 국내 코칭경력이 5년 이상인 전문가 6인을 대상으로 36개 코칭 행동에 대해 델파이 설문을 통해 내용 타당도를 점검하였다. 2차에 걸친 델파이 설문을 통해 최종적으로 35개의 코칭 행동이 선정 및 수정되었다. 35개의 코칭 행동들은 현장에서 얼마나 실천되고 있는지 Likert 5점 척도를 사용(5점 항상 실천함, 1점 전혀 실천하지 않음)하여 국내 코칭 전문가 20명에게 설문조사가 실시되었고, 대부분의 코칭 행동이 현장에서 잘 수행되고 있음이 확인되었다.

조대연, 박용호(2011)는 조대연, 김희영(2009)이 제안한 코칭 행동들 중에서 수행정도가 평균보다 높은, 즉 현장에서 상대적으로 빈번하게 행해지는 20개 코칭 행동 가운데 관리자의 코칭 행동과 부합하는 18개 문항을 선정하여 관리자 코칭 행동과 직무성과의 관계를 연구하였다. 조대연, 박용호(2011)은 관리자 코칭 행동으로 문제해결력 향상과 학습촉진을 제안했는데, 문제해결력 향상은 관리자와 구성원 간의 협력을 통한 성과향상을 위해 개인 또는 직무와 관련된 문제를 해결하는 데 초점을 두는 것을 의미하고, 학습촉진은 구성원에게 학습 기회를 제공하여 전문성 향상과 성장에 초점을 두는 것을 의미한다. 조대연, 박용호(2011)는 18개 문항에 대해 탐색적 요인분석과 Cronbach α 값을 바탕으로 조사도구의 타당도와 신뢰도를 검증하였다. 주성분 분석과 직각회전을 사용한 탐색적 요인분석은 요인부하량이 0.5 이상을 기준으로 4개의 문항이 제거되었고, 고유치가 1.0 이상을 충족하는 두 개의 요인이 최종적으로 도출되었다. 각 요인에 대한 신뢰도는 .927(문제해결력 향상), .910(학습촉진)으로 높은 신뢰도를 보였다. 조대연, 박용호(2011)에 의해 개발된 관리자 코칭 행동 측정도구는 최근까지도 조직 내 관리자 코칭 행동을 측정하기 위해 사용되고 있다. 예를 들어, 보험설계사를 대상으로 한 이윤수, 조대연, 김재민(2018)의 연구에서 확인적 요인분석 결과, 관리자 코칭 행동 18개 문항은 요인부하량 0.5 이상의 기준을 충족시켰고, Cronbach α 값도 0.8을 상회했다.

2. 조대연과 박용호(2011)의 관리자 코칭 행동 측정도구 활용 선행연구

조대연, 박용호(2011)의 측정도구를 사용한 실증연구 23편에서 요인분석 및 신뢰도 분석의 결과를 요약하면 다음과 같다.

<표 1> 조대연, 박용호(2011)의 관리자 코칭 행동 측정도구 활용 선행연구 검토

선행연구	표본	타당도 및 신뢰도 결과
강소운(2013)	항공사 종사자	문제해결력 향상 6문항($\alpha = .904$), 학습촉진 5문항($\alpha = .904$)

선행연구	표본	타당도 및 신뢰도 결과
김선우(2018)	유통업 종사자	문제해결력 향상 7문항($\alpha = .914$), 학습촉진 8문항($\alpha = .896$)
김세은(2018)	건설업 종사자	문제해결력 향상 4문항($\alpha = .927$), 학습촉진 3문항($\alpha = .886$)
김재민(2017)	손해보험설계사	문제해결력 향상 2문항($\alpha = .849$), 학습촉진 9문항($\alpha = .970$)
김지현(2012)	제조업 종사자	문제해결력 향상 6문항($\alpha = .922$), 학습촉진 5문항($\alpha = .915$)
김철영(2015)	일반 근로자	문제해결력 향상 4문항($\alpha = .872$), 학습촉진 7문항($\alpha = .916$)
김현근(2016)	대기업 종사자	문제해결력 향상 7문항($\alpha = .928$), 학습촉진 5문항($\alpha = .856$)
김홍준(2015)	출판업 종사자	문제해결력 향상 5문항($\alpha = .909$), 학습촉진 5문항($\alpha = .878$)
남기운(2017)	인수합병기업의 종사자	문제해결력 향상 7문항($\alpha = .961$), 학습촉진 6문항($\alpha = .943$)
문수미(2017)	제조업 종사자	문제해결력 향상 7문항($\alpha = .929$), 학습촉진 7문항($\alpha = .915$)
박주훈(2012)	대기업 종사자	문제해결력 향상 5문항($\alpha = .914$), 학습촉진 9문항($\alpha = .937$)
양유정(2016)	대기업 종사자	문제해결력 향상 4문항($\alpha = .909$), 학습촉진 6문항($\alpha = .900$)
이근원(2014)	중소기업 종사자	문제해결력 향상 7문항($\alpha = .907$), 학습촉진 8문항($\alpha = .813$)
이한중(2016)	손해보험사의 종사자	문제해결력 향상 5문항($\alpha = .868$), 학습촉진 4문항($\alpha = .874$)
이해정(2014)	대기업 종사자	문제해결력 향상 4문항($\alpha = .857$), 학습촉진 4문항($\alpha = .855$)
이희권(2014)	사무직 종사자	문제해결력 향상 5문항($\alpha = .898$), 학습촉진 6문항($\alpha = .911$)
장일우(2016)	손해보험 영업직	문제해결력 향상 6문항($\alpha = .944$), 학습촉진 9문항($\alpha = .966$)
전영화(2016)	한·중 대기업 종사자	문제해결력 향상 3문항($\alpha = .909$), 학습촉진 9문항($\alpha = .878$)
정용성(2017)	중국 현지채용인	문제해결력 향상 7문항($\alpha = .918$), 학습촉진 7문항($\alpha = .942$)
주두리(2016)	대기업 종사자	문제해결력 향상 7문항($\alpha = .889$), 학습촉진 6문항($\alpha = .885$)

<표 1>에 제시된 연구들은 요인부하량 0.5 기준에서 제거된 문항이 상이했고, 모두 주성분 분석(Principal Components Analysis, PCA)에 varimax 회전을 사용했다. 주성분 분석은 엄밀히 말해 요인분석이 아니라 자료의 축소에 적합한 방식이고, 많은 연구자들은 주성분 분석으로 요인분석을 수행해서는 안 되는 구체적인 이유와 문제점들을 지적해왔다(예를 들어, Fabrigar, Wegener, MacCallum, & Strahan, 1999). 요인 간 상관이 없도록 제약하는 직각회전(Orthogonal rotations)의 대표적인 회전방법인 varimax 회전 또한 관리자 코칭 행동과 같은 구인(construct)의 요인분석에는 적합하지 않고, 요인 간 상관을 허용하는 사각회전(Oblique rotations)이 더 권장된다(Fabrigar et al., 1999). 한편, 이지혜, 오상진(2017)의 연구에서 관리자 코칭 행동은 문제해결력 향상($\alpha = .874$)과 학습촉진($\alpha = .820$)은 높은 신뢰도를 보였고, 구조방정식을 활용하여 요인분석이 실시한 결과 적합도 지수가 권장 기준을 충족한 것으로 나타났다. 그러나 이는 학습촉진 요인 중 2개의 문항은 제거된 결과였다. 이윤수, 조대연, 김재민(2018) 연구에서는 문제

해결력 향상($\alpha = .864\sim.962$)과 학습촉진($\alpha = .888\sim.970$)에 대한 확인적 요인부석 결과 모든 문항이 타당한 것으로 분석되었다. 마지막으로 이지아(2015)의 연구에선 요인분석을 실시하지 않아 정보를 얻을 수 없었다.

이상의 결과를 종합해보면, 조대연과 박용호(2011)의 관리자 코칭 행동 측정도구는 이윤수, 조대연, 김재민(2018)의 연구를 제외하고는 온전히 2요인의 18문항 구조를 갖추고 있음이 실증적으로 확인된 바가 없었다. 조대연과 박용호(2011)의 연구에서도 18개 문항이 선정되었지만 실제 분석에선 총 14개 문항으로 문제해결력 향상($\alpha = .927$)과 학습촉진($\alpha = .910$)에서 각각 2문항씩 제거된 결과였다. 그럼에도 불구하고 이 측정도구는 신뢰도 계수가 .80 이상의 높은 값이 일관되게 보고된다는 점에서 기존의 요인분석 방식으로는 개선점을 모색하는 데 한계가 있어 보인다. 이에 본 연구는 문항반응이론의 관점에서 관리자 코칭 행동 검사도구를 재검토해 보고자 한다.

Ⅲ. 연구 방법

1. 자료 및 측정도구

본 연구는 이윤수, 조대연, 김재민(2018)에서 수집된 자료를 활용하였다. 왜냐하면 관리자 코칭 행동을 실증적으로 연구한 자료 중 가장 많은 숫자의 응답자 정보를 가지고 있고, 결측치가 없는 자료이며, 요인분석이 제대로 수행되었고, 2요인 18개 문항에서 문제가 발생하지 않았던 유일한 자료이기 때문이다. 이 자료는 관리자와 조직원의 응답으로 구분되어 있는데, 본 연구에서는 847명의 조직원이 응답한 데이터만 활용되었다. 손해보험 설계사를 대상으로 수집된 자료는 보험사의 특성이 반영되어 여성(737명, 87.0%)이 남성보다 많았고, 평균 연령은 약 49세였다. 총 경력은 평균 14년이었고, 학력은 과반수가 중졸(56.8%)로 대기업에 비해 낮은 편이었다(보다 구체적인 정보는 이윤수, 조대연, 김재민(2018)의 연구 참고). 측정도구인 조대연과 박용호(2011)의 관리자 코칭 행동은 문제해결력 향상과 학습촉진 2개 요인이고, 문항 수는 요인당 9문항씩 총 18개 문항이다. 응답 범주는 전혀 그렇지 않다(1점)에서 매우 그렇다(5점)로 구성된 Likert 5점 척도가 사용되었다. 이윤수, 조대연, 김재민(2018)의 연구에서 확인적 요인분석 결과, 모든 문항은 요인부하량 .50 이상의 값을 나타냈고, 신뢰도 또한 .90을 상회했다(문제해결력 향상 .962, 학습촉진 .970). 모델 적합도 지수는 $\chi^2(103) = 651.701$, TLI = .963, CFI = .968, RMSEA = .079 [.074, .085], SRMR = .019로 적합도 지수 기준(Browne & Cudeck, 1993; Hu & Bentler, 1999)에 따라 양호한 것으로 나타났다.

2. 문항반응이론 모형

문항반응이론(Item Response Theory, IRT) 모형은 검사 총점에 의존하고 진점수(true score)를 추정하는 고전검사이론(Classical Test Theory; CTT)과 달리 검사도구로 측정된 심리적 특질(trait)과 문항 응답 간의 관계를 보여준다. 특질이란 검사도구로 추정되는 응답자의 점수(IRT score)로 능력(ability)이라고도 불린다. 본 연구는 Likert 척도를 사용하여 태도나 상태를 설문하는 경우로 능력보다는 특질이라는 용어를 사용하기로 한다. Rasch 모형은 문항의 적합도, 문항의 곤란도, 응답 범주(item category)의 적절성을 평가할 수 있다는 장점이 있다(이익섭, 홍세희, 신은경, 2007). 문항반응이론은 검사도구의 개발에도 사용될 수 있다. 본 연구를 예로 들면, 관리자 코칭 행동을 측정하는 데 있어 적합도가 떨어지는 문항을 선별해주고, 응답자가 동의하기 쉬운 수준부터 어려운 수준까지 고르게 질문하고 있는지 검토하며, 현재 5점 척도로 임의로 결정된 응답 범주의 수와 간격이 적절한지 평가할 수 있다.

문항의 적합도(MNSQ)는 비가중 평균자승(unweighted mean square; Outfit)과 가중 평균자승(weighted mean square; Infit) 지수로 판단하는데, 1에 가까울수록 좋고 1.3 이상(해당 문항이 다른 문항과 지나치게 동떨어져 있음을 의미)이나 0.7 이하(해당 문항이 다른 문항과 지나치게 중복됨을 의미)는 좋지 않은 적합도를 나타낸다(Bond & Fox, 2015). 예를 들어, 어떤 측정도구는 유사해 보이는 여러 문항들로 구성된 경우가 있는데, 이러한 측정도구는 응답의 일관성이 높기 때문에 대체로 신뢰도가 높게 나타나고, 문항이 구인을 잘 측정하고 있다면 다른 문항들과 중복이 되더라도 타당도에서 문제가 없을 가능성이 높다. 하지만 Rasch 모형은 다른 문항과 지나치게 중복되는 문항을 식별해주기 때문에 측정도구를 개발하는 과정에서 유사한 문항들을 줄여나가는 과정에서 활용될 수 있다.

문항 곤란도(item difficulty) 또한 대상 모집단의 특성 수준에 맞는 문항을 개발하는 데 유용하다(DeMars, 2010). 예를 들어, 대학생의 학업 소진을 측정하기 위해서 직장인의 직무 소진 측정도구를 적용한다면 모집단의 특성이 달라지기 때문에 해당 구인을 제대로 측정할 수 없을 것이다. 두 집단에서 소진을 구성하는 요인 자체가 달라질 수 있고, 대부분의 대학생은 직장생활을 경험하지 않았기 때문에 일부 문항은 아예 이해할 수 없거나 동의하기 곤란할 수 있다. Rasch 모형에서 문항 곤란도는 로짓(logit)으로 변환되어 개인의 특질 점수와 직접적으로 비교됨으로써 곤란도의 분포가 개인 특질의 모든 범위를 측정하고 있는지를 점검할 수 있다(이익섭, 홍세희, 신은경, 2007; 홍세희 외, 2018). 위의 예에서 대학생들은 직장인의 직무 소진을 제대로 이해할 수 없기 때문에 문항에 동의하기 어려울 수 있기 때문에 Likert 5점 척도에서 ‘전혀 동의하지 않는다’ 나 ‘동의하지 않는다’를 선택할 가능성이 높다. 결과적으로 곤란도의 분포는 한쪽으로 쏠리게 되는데, Rasch 모형은 이를 시각적으로 점검할 수 있다. jMetrik

통계 프로그램에서 infit과 outfit, 문항 곤란도 분석 결과는 Rasch models을 선택하면 확인할 수 있고 곤란도의 분포는 item map을 선택하면 확인할 수 있다.

마지막으로 Rasch 모형은 응답 범주의 수와 간격의 적절성을 평가할 수 있는 장점이 있다. 응답 범주의 적절성은 개인의 특질을 x축으로 하고, 각 응답 범주가 선택될 확률을 y축으로 하는 범주 확률 곡선(category probability curve)을 통해 평가된다(홍세희 외, 2018). jMetrik 통계 프로그램에서 범주 확률 곡선은 IRT plot을 선택하면 확인할 수 있다. 위의 예에서 문항 곤란도 문제가 해결되었다고 가정하더라도 대학생들이 ‘매우 동의한다’라고 응답을 아무도 선택하지 않는다면 해당 선택지를 제거하는 것을 고려할 수 있다. 이처럼 Rasch 모형은 문항 및 범주를 수정 및 제거하는 데 활용될 수 있는데, 그동안 주로 확인적 요인분석(confirmatory factor analysis)을 통해 타당도를 검증한 척도에 완성도를 높여주는 데 활용되어 왔다. 예를 들어, 홍세희와 조용래(2006)는 70문항으로 구성된 역기능적 신념검사도구에 Rasch 모형을 적용하여 36문항으로 구성된 단축판 검사도구를 개발하였고, 응답 범주를 기존의 7점 척도에서 6점 척도로 수정한 바 있다.

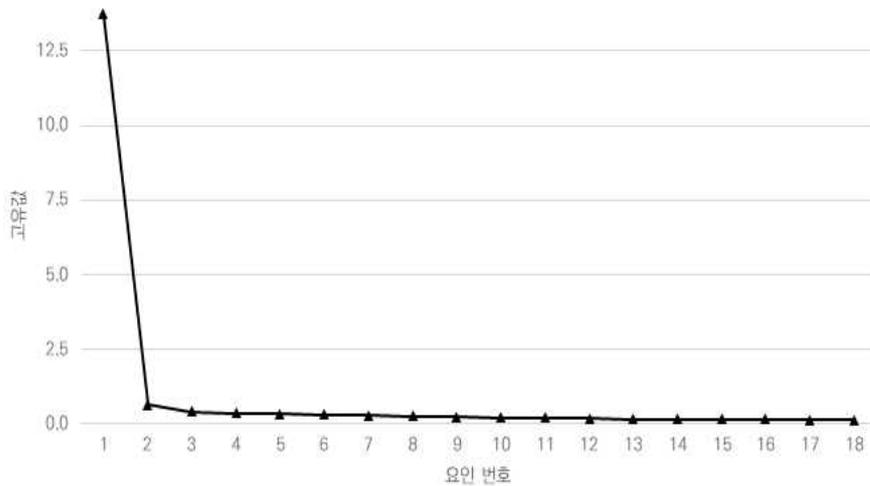
Rasch 모형(Rasch, 1960)은 문항반응이론 모형 중 이분화된 반응 모형(Dichotomous Response Model)이고, 이것을 다점척도로 확장한 부분점수모형(Partial Credit Model; Master, 1982)과 평정척도모형(Rating Scale Model; Andrich, 1978)이 있다. 본 연구에서 관리자 코칭 행동 척도는 다점척도 모형(polytomous model) 중 Andrich(1978)의 평정척도모형을 적용하여 문항의 적합도, 곤란도, 응답 범주의 적절성을 평가함으로써 측정도구를 개선하고자 한다. 평정척도모형은 문항의 곤란도는 다르지만 응답구조는 문항에 따라 동일하다고 가정하기 때문에 응답 범주의 간격이 동일하다고 간주되는 Likert 척도를 사용한 경우 평정척도모형이 일반적으로 사용된다(Hong, Kim, & Wolfe, 2005). 본 연구에서 평가할 관리자 코칭 행동 척도 또한 Likert 5점 척도를 사용하여 측정도록 고안되었기 때문에 평정척도모형을 사용하는 것을 선택했다. jMetrik 통계 프로그램에서 평정척도모형의 선택은 group을 동일한 문자로 제약하는 방식으로 적용할 수 있다. 분석에는 SPSS와 jMetrik 통계 프로그램이 사용되었다.

IV. 연구 결과

1. 일차원성 검증

분석의 첫 번째 단계로 일차원성(Unidimensionality)를 확인하기 위해 스크리 도표를 활용하였다. 일차원성은 문항반응이론의 가정 중 하나로, 관리자 코칭 행동은 2개의 요인으로 구분되지만 그 요인들이 관리자 코칭이라는 한 가지 개념을 측정하는지 평가하는 것을 의미한다. 이

는 스크리 도표로 확인이 가능하며, 추정방법으로 SPSS의 주축분해법이 적용되었다. [그림 1]에서 알 수 있듯이, 하나의 특질이 대부분의 총분산(68.81%)을 설명하기 때문에 일차원성이 성립되는 것으로 해석된다(DeMars, 2010).



[그림 1] 스크리 도표

2. 문항 분석 결과

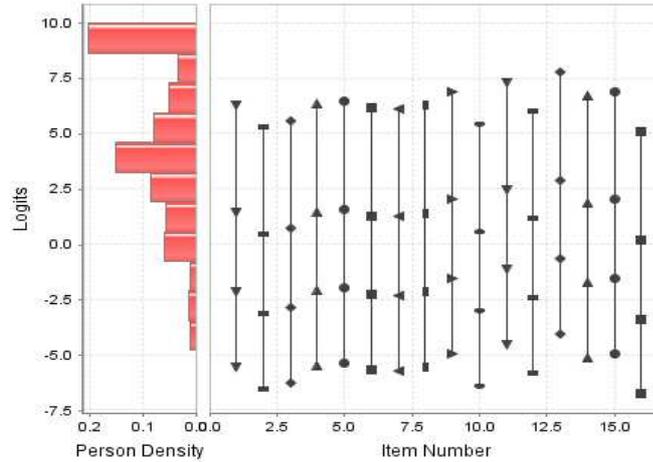
다음으로 문항 곤란도(item difficulty), 피험자 적합도 분석(MNSQ), 문항 및 피험자 신뢰도를 분석하였다. 분석 결과, 문항 곤란도는 -0.64 (평균 4.31)에서 0.74 (평균 4.08)로 다양하게 분포되었으나 일부 문항들은 유사한 곤란도를 가졌다(표 2) 참고). 예를 들어, 11번, 16번, 17번 문항은 같은 요인에 속하면서 곤란도는 각각 0.29, 0.31, 0.36으로 차이가 크지 않았다. 적합도 지수 분석 결과, 1번 문항과 2번 문항은 outfit과 infit이 각각 1.4이상으로 분석되어 제거할 필요성이 있었다. 이는 해당 문항이 다른 문항들과 다소 동떨어져 있음을 의미하는데, 1번 문항은 업무수행과 관련한 문제를 다룬다는 점에서 다른 문항들과 유사했지만 상황을 진단하고, 계획수립을 돕는 등 구체적인 문제해결력 향상과의 관련성은 상대적으로 떨어지는 것으로 판단된다. 2번 문항 또한 다른 문항과 달리 업무수행 과정에서 직면할 수 있는 문제 상황과는 거리가 멀고, 상사와의 신뢰, 우호적 관계를 측정하는 문항에 더 가까워보인다. 두 문항을 제거한 뒤 문항신뢰도는 .912, 피험자신뢰도는 .960으로 해석 기준 .80 이상(Fox & Jones, 1998)을 충족했다. 이러한 결과는 관리자 코칭 행동 측정도구를 사용한 선행 연구에서도 .90 이상의 높은 신뢰도가 일관되게 보고된 것과 일치하는 결과이다.

<표 2> 문항 분석 결과

문항	평균	곤란도	MNSQ	
			infit	outfit
1. 나의 상사는 나의 업무수행과 관련한 당면한 문제 또는 이슈 등을 파악한다.	4.31	-0.64	1.50	1.72
2. 나의 상사는 상호신뢰를 기반으로 나와 인간적인 관계를 형성한다.	4.26	-0.33	1.46	1.51
3. 나의 상사는 나의 업무수행 관련 어려움을 파악하기 위해 나의 현재 상황을 진단한다.	4.20	0.07	1.13	1.11
4. 나의 상사는 내가 업무수행 관련 어려움을 극복할 수 있는 방법을 찾도록 도와준다.	4.27	-0.39	0.97	0.98
5. 나의 상사는 내가 업무수행의 바람직한 상태를 결정할 수 있도록 도와준다.	4.25	-0.27	0.91	0.92
6. 나의 상사는 나의 업무수행 향상을 위해 구체적 목표를 설정할 수 있도록 함께 논의하고 결정한다.	4.19	0.08	0.89	0.86
7. 나의 상사는 나의 업무수행 향상을 위한 계획수립을 돕고, 필요하다면 계획수정을 지원한다.	4.19	0.14	0.91	0.87
8. 나의 상사는 나의 업무수행 향상 계획을 실천할 수 있도록 나에게 주도권을 준다.	4.21	0.00	1.14	1.23
9. 나의 상사는 나에게 도전의식을 높일 수 있는 조언과 지원을 제공한다.	4.21	-0.01	0.85	0.79
10. 나의 상사는 나의 역량개발을 위해 지속적으로 학습기회를 발전시킬 수 있도록 돕는다.	4.20	0.05	0.90	0.82
11. 나의 상사는 건설적인 피드백을 제공하여 나의 (업무수행)행동을 강화한다.	4.15	0.34	0.82	0.80
12. 나의 상사는 중요한 정보를 제공할 때 나의 주의를 집중시킨다.	4.26	-0.33	0.92	0.95
13. 나의 상사는 타인의 입장이 되어 볼 수 있는 기회를 통해 나의 가치관(관점)을 넓혀준다.	4.12	0.54	0.96	0.97
14. 나의 상사는 내가 학습한 것과 현재 상황이 연계될 수 있도록 돕는다.	4.22	-0.06	0.86	0.84
15. 나의 상사는 나와 함께 수립한 계획의 실천과정을 정기적으로 점검하고 지원한다.	4.08	0.74	0.86	0.85
16. 나의 상사는 나에게 목표달성을 위해 학습한 것을 실천해 볼 수 있도록 돕는다.	4.17	0.26	0.82	0.77
17. 나의 상사는 내가 업무수행 향상을 위해 노력하는 동안 발생할 수 있는 어려움과 장애물을 파악할 수 있도록 돕는다.	4.15	0.34	0.89	0.87
18. 나의 상사는 나의 성취와 능력함양을 칭찬하고 격려해준다.	4.29	-0.51	1.09	1.02

다음으로 표준화된 잔차요인분석(standardized residual factor analysis, Linacre, 1998)을 실시하였다. 측정하고자 하는 개념이 설명된 후 잔차에서 고유값(Eigen value) 기준으로 두드러지게 나타나는 요인이 있는지 분석하는 것으로, 두드러지게 나타나는 요인이 없다면 일차원성을 만족하고, 문항 적합도에도 문제가 없음을 의미한다. 분석 결과, 그러한 요인은 없는 것으로 나타났다(F1: 2.20, F2: 1.62, F3: 1.33, F4: 1.28, F5: 1.20).

마지막으로 개인의 특질 점수와 문항의 곤란도를 로짓(logit)으로 변환하면 동일한 척도를 갖기 때문에 [그림 2]와 같이 직접적인 비교가 가능하다. 관리자 코칭 행동과 같이 Likert 다점 척도로 구성된 검사도구에서 문항의 평균은 동의하기 쉬운 정도인 문항 곤란도(난이도)를 의미하고, 평균이 높을수록 동의하기 쉬운 문항을 뜻한다. 두 분포가 비슷하면 문항의 곤란도가 개인의 특질의 범위를 적절하게 측정할 수 있다고 해석된다(Hong et al., 2005). [그림 2]의 왼쪽 분포에서 알 수 있듯이, 개인의 특질 점수는 정상분포처럼 보이다가 아주 높은 특질을 가진 사람들에 몰려있는 분포를 띤다. 이는 대부분의 사람들이 관리자 코칭 행동을 높게 응답했음을 의미한다. 실제로 각 문항의 평균 점수는 표1에서 알 수 있듯이 모두 4점 이상으로 이러한 결과를 지지한다. 한편, 오른쪽 분포에서 문항 곤란도는 아주 높은 특질을 가진 사람들에게는 적절하지 않다. 즉, 현재의 검사도구는 관리자 코칭 행동을 어느 정도 수준 이상으로 인식하는 사람들에게 지나치게 쉽게 동의할 수밖에 없는 문항들로 고안되어 있음을 의미한다. 이는 이 검사도구가 최초 현장에서 발휘되는 빈도가 높은 문항을 선별하여 구성되었다는 점에서 이해가 되는 결과이지만 동시에 관리자 코칭 행동이 비교적 높은 사람들에게 적합한 난이도와 그들을 변별할 수 있게끔 개선되어야 함을 시사한다. 그렇지 않을 경우, 응답자는 관리자 코칭 행동을 일정 수준 이상으로만 인식한다면 대부분의 문항에 전적으로 동의한다는 하나의 응답만 선택할 가능성이 높아지고, 통계분석 결과 또한 왜곡 및 편향될 가능성이 있다.

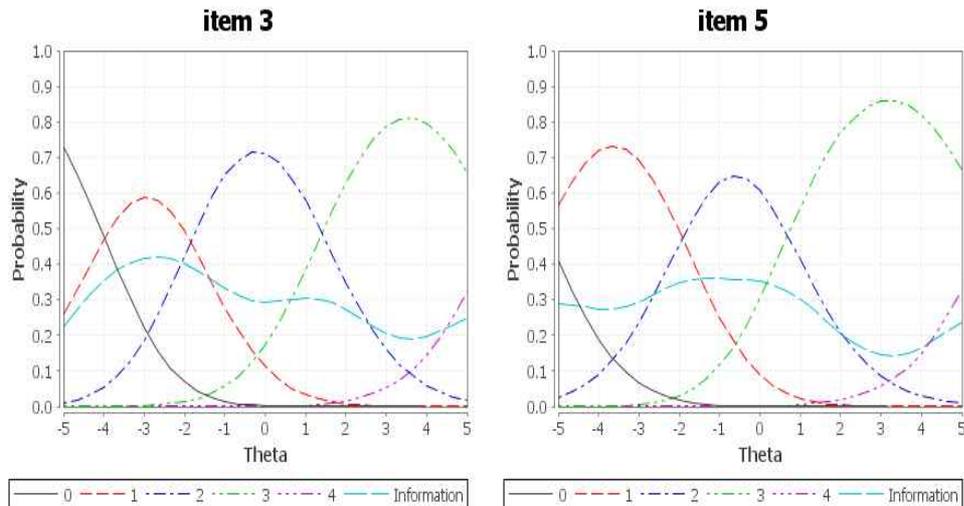


[그림 2] 개인 특질 점수와 문항 곤란도 분포 비교

3. 응답 범주의 타당성 검증

다음으로 Likert 5점 척도로 구성된 관리자 코칭 행동 측정도구의 응답 범주가 타당한지 검토하였다. 우선 각 응답 범주가 증가할수록 Threshold 값이 증가하는지, 각 범주에 대한 적합도(infit, outfit) 기준을 만족하는지 검토했다(Hong & Wong, 2005). 검토 결과 응답 범주가 증가할수록 Threshold 값이 증가(-5.55, -2.15, 1.42, 6.29)했고, 적합도 기준도 충족되었다. 그러나 개인의 특질, 즉 개인이 보유한 관리자 코칭 행동의 수준을 x 축으로 하고, 1점(“전혀 동의하지 않는다”)부터 5점(“전적으로 동의한다”) 선택지가 선택될 확률을 y 축으로 하는 범주 확률 곡선 또한 5점 척도로 구성된 응답 범주에 문제가 있음을 암시했다. [그림 3]은 관리자 코칭 행동 문항 중 3번 문항과 5번 문항에 대한 범주 확률 곡선을 보여준다. 0번 곡선부터 4번 곡선은 각각 1번부터 5번 선택지의 확률 곡선을 나타낸다. 즉, 3번 문항의 경우, 개인의 특질(θ)이 -4 이하일 경우(관리자 코칭 행동을 아주 낮게 인식한다는 의미) 1번 선택지를 선택할 확률이 높고, θ 가 -4와 -2 사이인 경우 2번 선택지를 선택할 확률이 높음을 의미한다. 한편, 4번 곡선(5번 선택지)은 그림에서 절단되었지만 절단된 오른쪽에 존재하는데, 이는 대부분의 사람들이 관리자 코칭 행동을 아주 높게 응답했음(θ 가 아주 높음)을 암시한다. 그런데 3번 문항(9, 10, 18번 문항도 비슷한 분포)의 왼쪽의 경우 1번 선택지가 선택될 확률이 있으나, 5번 문항(4, 5, 6, 7, 8, 11, 12, 13, 14, 15, 16, 17번 문항도 비슷한 분포)에는 1번 선택지가 선택될 확률이 거의 없었다. 실제로 모든 문항에서 1번 선택지(전혀 그렇지 않다)로 응답한 응답자는 1% 미만으로 극소수였다. 이는 응답 범주의 조정이 필요함을 암시하는데, 여러 가지 방

법이 고려해볼 수 있다. 첫 번째 방법은 1번 선택지가 전혀 선택되지 않았다는 점에서 1번 선택지를 제거하거나 2번 선택지와 결합하는 방식을 고려해볼 수 있다. 빈도분석 결과 1번 선택지 “전혀 동의하지 않는다”가 거의 선택되지 않았다는 점에 착안한 방법이다. 두 번째 방법은 문항의 곤란도를 조정하여 개인 특질이 높은 사람들에게 적정 수준의 난이도로 만드는 것이다. [그림 3]에서 알 수 있듯이, 현재 검사도구는 개인 특질이 높은 사람들에게 너무 쉽게 구성되어 있고, 이는 1번 선택지를 선택한 사람이 극히 드물게 된 하나의 원인일 수 있다. 그러나 현재 자료로는 두 번째 방법을 검증할 수 없다. 같은 이유로 세 번째 방법인 척도를 늘리는 방법(예를 들어, 7점 척도)도 현재 자료로는 검증할 수 없지만 검사도구를 개선하기 위한 방안으로 고려될 수 있다. 척도를 세분화하면 개인 특질이 높은 사람들 간의 미세한 차이를 구분할 수 있어 관리자 코칭 행동을 높게 응답한 사람들을 변별하는 데 기여할 수 있기 때문이다. 그러나 현재 자료의 분포로 미루어 짐작해보면 개인 특질이 낮은 사람들 중에서 특정 선택지를 선택하지 않을 가능성(1번 선택지를 선택하지 않을 가능성)이 재발할 수 있다.



[그림3] 응답 범주 확률커브

4. 검사도구의 요인 분석

마지막으로 검사도구 개발 관점에서 적합도 기준을 충족하지 못한 2개의 문항을 제외한 후

탐색적 요인분석을 실시했고, 조대연과 박용호(2011)의 차원 구분에 따라 확인적 요인분석도 실시했다. 일반적으로 동일한 표본으로 탐색적 요인분석과 확인적 요인분석을 실시하는 것은 지양된다. 그러나 Rasch 모형으로 분석한 결과 2개의 문항만 제거되는 변화가 있었고, 스크리도표 분석 결과 관리자 코칭 행동은 1요인으로 확연하게 구분되는 것으로 보였기 때문에 두 가지 방식을 모두 활용하였다. 먼저 탐색적 요인분석은 1요인 모형부터 3요인 모형을 대안모형으로 개발하여 비교하였다. 분석 결과는 다음 <표 3>과 같다.

<표 3> 적합도 지수 분석 결과

적합도 지수	χ^2	df	TLI	CFI	RMSEA	SRMR
기준값	-	-	.9이상	.9이상	.08이하	
1-factor model	857.585***	104	.949	.956	.092 [.087, .098]	.021
2-factor model	437.730***	89	.973	.980	.068 [.062, .074]	.013
3-factor model	284.871***	75	.980	.988	.057 [.050, .065]	.010
2-factor CFA model	651.701***	103	.963	.968	.079 [.074, .085]	.019

적합도 지수를 기준으로 3모형이 더 적합해보였으나 실제로 마지막 요인에서 요인부하량 .50을 넘는 문항은 하나도 없었다. 2요인 모형의 경우 대체로 잘 구분되었으나 7, 8, 9번 문항에서 요인부하량 기준을 충족하지 못했거나 다른 요인에 로드되었다. 반면, 1요인 모형에서는 모든 문항이 기준을 충족했다. 한편, 확인적 요인분석 결과, 2요인 모형에서 모든 문항들이 적합도 기준을 충족했다(<표4> 참고). 그러나 두 요인 간 상관인 .965로 다중공선성이 의심되었다.

<표 4> 2요인 모형 탐색적·확인적 요인분석 결과

문항	탐색적 요인분석 요인부하량		확인적 요인분석 요인부하량	
	문제해결력 향상	학습촉진	문제해결력 향상	학습촉진
문제해결력 향상 3	.678	.199	.857	
문제해결력 향상 4	.883	.021	.879	
문제해결력 향상 5	.935	-.022	.888	
문제해결력 향상 6	.707	.212	.905	
문제해결력 향상 7	.468	.439	.894	
문제해결력 향상 8	.132	.705	.821	
문제해결력 향상 9	.347	.565	.892	
학습촉진 10	.266	.642		.892
학습촉진 11	.299	.620		.899
학습촉진 12	.088	.786		.866
학습촉진 13	.005	.882		.883
학습촉진 14	-.064	.954		.890
학습촉진 15	-.019	.927		.906
학습촉진 16	-.075	.979		.905

문항	탐색적 요인분석 요인부하량		확인적 요인분석 요인부하량	
	문제해결력 향상	학습촉진	문제해결력 향상	학습촉진
학습촉진 17	.078	.817		.889
학습촉진 18	.120	.736		.846

이상의 결과를 종합하면 관리자 코칭 행동은 2개의 하위요인으로 구성되기보다는 하나의 요인으로 보는 것이 적절할 수 있음을 시사한다. [그림 1]의 스크리 도표 분석 결과도 1요인을 지지한다.

V. 논의 및 결론

연구 결과를 바탕으로 관리자 코칭 행동 검사도구의 개선점과 그 결과를 바탕으로 인적자원 개발 분야에 시사하는 바를 논의하면 다음과 같다. 첫째, 관리자 코칭 행동 검사도구의 문항 곤란도는 조정되어야 한다. 개인 특질 점수와 문항 곤란도 분포 비교해보면, 개인의 특질 점수가 높은 사람에게 적합한 문항 곤란도를 가진 문항은 없었다. 다시 말해, 관리자 코칭 행동에 대한 인식이 극단적으로 높지 않은 사람들에게 현재 검사도구의 문항들은 동의하기 지나치게 쉬운 것으로 나타났다. 각 문항에 대한 평균도 4.0점을 상회했고, 본 연구뿐만 아니라 선행 연구에서도 신뢰도 지수가 일관되게 높게 보고된 것은 문항이 동의하기 쉽게 구성된 결과임을 암시한다. 물론 관리자 코칭 행동에 대한 인식이 실제로 높을 수 있다. 하지만 관리자 코칭 행동 검사도구는 관리자에 대한 조직원의 평가를 상정하기 때문에 익명성이 보장되는 설문조사임에도 불구하고 조직원들이 긍정적으로 평가하는 경향이 있을 수 있음을 감안해야 한다. 관리자가 인식하는 코칭 행동과 부하직원이 인식하는 코칭 행동을 비교해봄으로써 이를 검증할 수 있다. 이윤수, 정희원, 조대연(2016)은 행위자-상대자 상호의존 모형을 활용하여 리더가 생각하는 공유 리더십과 구성원이 생각하는 공유 리더십에 간극이 있을 수 있음을 실증적으로 보여준 바가 있었는데, 이 방법론은 관리자가 인식하는 자신의 코칭 행동과 부하직원이 인식하는 관리자의 코칭 행동에 차이가 있는지, 그리고 그 차이가 어떤 다른 결과를 낳는지에 대한 일종의 해답을 줄 수 있다.

다른 측면에서 현재의 관리자 코칭 행동 검사도구가 동의하기 쉽게 구성된 원인을 검사도구 개발 과정에서 찾을 수 있다. 조대연, 박용호(2011)는 문항을 선정할 때, 코칭 전문가들에 의해서 현장에서 얼마나 자주 실천되고 있는지 그 빈도에 대한 평균을 기준으로 삼았다. 따라서 각 문항들이 담고 있는 코칭 행동들은 현장에서 자주 접할 수 있는 것들이다. 게다가 각 문항들을 살펴보면, 관리자 코칭 행동이 평이하게 서술되어 있는 형식을 취하는데, 동일한 코칭 행동을 측정하는 문항이라고 하더라도 곤란도를 높이면 관리자 코칭 행동에 대한 인식이 높은

사람들에게 적절한 곤란도를 갖는 문항이 될 수 있다. 예를 들어, 3번 문항은 “나의 상사는 나의 업무수행 관련 어려움을 파악하기 위해 나의 현재 상황을 진단한다” 로, “나의 상사는 급한 업무가 있더라도...” 혹은 “나의 상사는 매주...” 등과 같이 관리자의 코칭 행동의 상대적 우선순위, 주기, 빈도의 조건을 추가하는 식으로 곤란도를 높일 수 있을 것이다. 또 다른 예로 16번 문항은 “나의 상사는 나에게 목표달성을 위해 학습한 것을 실천해 볼 수 있도록 돕는다” 로, “학습한 것을 현장에서 직접 실천해 볼 수 있도록 동행하여 돕는다” 등과 같이 보다 구체적인 코칭 행동을 제시하는 식으로 수정될 수 있을 것이다. 문항 곤란도를 조정하는 것은 모든 문항에 해당하는 것은 아니고, 전체 검사도구의 곤란도가 고르게 분포될 수 있게 하는 것이 바람직하다. 예를 들어, 문항 곤란도가 유사했던 문항들(예를 들어 11번 문항 0.29, 16번 문항 0.31, 17번 문항 0.36) 중 일부 문항의 곤란도를 조정하는 것이 효율적이다. 또는 조대연과 박용호(2011)에 의해 선택되지 않은 문항들을 재검토해서 추가하거나 기존 문항에 결합하는 방법도 고려될 수 있다. 그 문항들은 현장에서 발생 빈도는 낮지만 일부는 관리자 코칭 행동이 아주 높을 때에만 발휘되는 것일 수 있고, 이 경우 문항 곤란도가 높으면서 내용적으로 타당한 문항 개발의 재료로서의 가치가 있기 때문이다.

둘째, 관리자 코칭 행동 하위요인과 일부 문항의 적합성에 대해 추가적인 논의되어야 한다. 관리자 코칭 행동의 두 가지 하위 요인은 두 가지 관리자 기능적 구성요소에 이론적 기반을 두고 개발되었다. 본 연구에서도 확인적 요인분석 결과 관리자 코칭 행동이 두 가지 하위 요인으로 구성되는 것이 지지되었다. 하지만 두 요인은 상관계수가 0.9 이상으로 실제로 구분되기 어려운 변수로 판단되고, 이는 회귀분석 등 통계분석에서 다중공선성을 유발할 수 있다. 스크리 도표의 패턴도 관리자 코칭 행동이 하나의 요인으로 간주해도 무방할 정도로 일차원성을 지지하는 결과가 나왔다. 선행연구에서는 요인분석의 방법이 제대로 적용되지 않았거나 많은 문항이 제거되었다는 점을 감안해야 하지만 22개 연구 중 21개 연구에서 두 변인의 상관이 0.7을 상회했고, 일부 연구(김지현, 2012, $r=.829$; 김현근, 2016, $r=.825$; 남기운, 2017, $r=.845$; 문수미, 2017, $r=.826$; 박주훈, 2012, $r=.831$; 양유정, 2018, $r=.801$; 이해정, 2014, $r=.831$; 장일우, 2016, $r=.885$)에서는 0.8을 상회하는 등 매우 높게 분석되었다. 실무적으로도 관리자 코칭 행동이 두 가지 하위요인이 구분되어 실천되기보다는 관리자는 학습촉진을 통해서 문제해결력을 향상시킬 수 있고, 문제해결력을 향상시키기 위해 학습촉진을 제공할 수 있다. 따라서 보다 유연하고 종합적으로 관리자 코칭 행동을 바라보는 관점을 채택할 필요성에 대해 추가적인 논의가 필요해 보인다. 이 논의가 중요한 까닭은 관리자 코칭 행동의 교육 훈련에 있어서 두 가지 기능을 개별적 또는 통합적으로 접근하는 것과 관련되기 때문이다. 보다 궁극적이고 확장적으로 조대연, 박용호(2011)의 기능적 관점을 떠나 강영순, 김정훈(2013), 조은현(2010)이 제안한 구성요소들과의 이론적 결합이나 통합에 대한 추가적인 연구도 필요할 것으로 보인다.

셋째, 문항과 관련하여 적합도 분석 결과, 2개의 문항이 제거될 필요가 있는 것으로 나타났

다. 1번 문항(“나의 상사는 나의 업무수행과 관련한 당면한 문제 또는 이슈 등을 파악한다”)은 상술했듯이 구체적인 문제해결력 향상을 위한 코칭 행동을 담고 있지 못하는 것으로 보이고, 2번 문항(“나의 상사는 상호신뢰를 기반으로 나와 인간적인 관계를 형성한다”)은 문제해결력 향상과는 거리가 멀고, 상사와의 신뢰적 관계를 측정하는 문항에 더 가까워보인다. 그러나 본 연구는 두 문항이 반드시 제거되어야 함을 주장하기 보다는 문항의 전체적인 곤란도를 조정하는 과정에서 해당 문항이 수정될 수 있고, 그 뒤 요인분석뿐만 아니라 문항반응이론에 기반을 두고 다각도로 문항들을 검토해야 할 필요성이 있음을 강조하는 데 무게를 둔다.

셋째, 응답 범주를 변경하는 것을 고려해야 한다. 관리자 코칭 행동뿐만 아니라 대부분의 검사도구의 응답 범주는 임의적으로 정해진다. 물론 Likert 5점 척도는 인적자원개발 분야의 설문조사에서 가장 빈번하게 활용되는 응답 범주로 설문응답자에게 친숙하다는 장점이 있다. 하지만 본 연구에서 드러났듯이, 대부분의 응답자가 4점 또는 5점을 선택했는데, 문항 곤란도 조정을 통해서 4점과 5점에 대한 쓸림 현상은 어느 정도 해소될 것으로 보이나 분포가 정상분포에 가깝게 변화될 것인지에 대해서는 예단할 수 없다. 예를 들어, 곤란도 조정에 성공해 기존에 n점을 선택한 사람들 중 일부가 n-1점을 선택하면서 분포가 이동할 수는 있지만, 분포의 모양은 변하지 않아 여전히 많은 사람들이 5점을 선택하는 결과가 나올 수 있다. 본 연구에서는 척도를 늘리거나 일부 척도를 결합하는 방식이 있을 수 있음을 제안했으나 어느 방식이든 문항 곤란도 수정이 선행된 후 재검토가 이루어져야 하고, 관리자 코칭 행동 검사도구에 국한되지 않고 검사도구의 개발에 있어서 응답 범주에 대한 보다 세심한 고려가 필요함을 강조하고자 한다.

이처럼 문항반응이론에 입각한 문항 곤란도, 문항의 적절성, 응답 범주의 적절성에 대한 고려는 인적자원개발 분야에서 검사도구를 개발할 때 있어 이론적·실천적 시사점을 제공해준다. 인적자원개발 분야의 양적 연구는 타 분야에서 개발되고 검증된 검사도구를 쓰는 경우도 있지만 관리자 코칭 행동과 같이 자체적으로 검사도구를 개발하는 경우도 있다. 경우에 따라서는 기존의 검사도구를 연구 목적에 부합하는 형태로 수정할 필요도 있다. 예를 들어, 리더십의 주체를 교사에서 상사로 바꾸거나 작업환경을 학교에서 직장으로 바꾸는 등 모집단의 특성이 달라지는 경우, 유사하게 보이는 문항이더라도, 요인분석이나 신뢰도 분석 결과도 기준을 충족시키더라도, 문항 곤란도나 적절성은 달라질 수 있다. 그러나 인적자원개발 분야의 국내 학술지 논문들(예를 들어, 평생교육·HRD연구, 한국HRD연구, HRD연구, 역량개발학습연구 등) 중에서 문항반응이론이 적용된 사례는 2019년 2월 현재 없다. 응답 범주에 대한 고려도 보다 객관적이고 정교한 검사도구 개발을 위해 고려되어야 한다. 국내에서 개발된 대부분의 측정도구가 그러하듯 관리자 코칭 행동 측정도구의 응답 범주는 특별한 이유 없이 Likert 5점 척도로 측정되었다. 극단적인 예로 직무열의(work engagement)는 Likert 7점 척도로 측정되도록 고안되었는데, 국내 연구에서는 임의로 Likert 5점 척도로 측정되는 경우를 쉽게 찾아볼 수 있다.

이는 인적자원개발 분야의 연구자에게만 해당하는 문제는 아니며, 전술했듯이 응답 범주 설정에 대한 이론적·합리적 의사결정이 적극 권장되어야 할 것이다. 보다 정교하고 객관적인 검사도구의 개발은 인적자원개발 분야의 연구의 질과 연구 결과의 신빙성을 담보해줄 것이다. 실천적인 관점에서도 정교하고 객관적인 측정도구의 개발은 관리자의 코칭 행동을 적확하게 측정하는 데 필요하다. 관리자의 코칭 행동은 관리자의 성과 평가에 반영될 수도 있고, 관리자의 리더십을 대변하는 척도로 활용될 수도 있으며, 관리자의 교육훈련 내용으로 포함될 수도 있다. 그러나 관리자 코칭 행동을 제대로 측정할 수 없다면 그 측정된 결과를 성과에 반영하는 것에 합의를 이끌어내기 어렵고, 양질의 교육훈련을 고안해낼 수도 없을 것이다.

이상의 결과를 바탕으로 후속연구를 위한 제언으로 본 연구에서 제기된 문제점들을 고려하여 후속연구에서는 문항의 난이도를 조정하고, 제거가 필요한 문항을 수정하며, 응답 범주에 대해 고민해볼 수 있다. 그리고 다양한 산업, 직업, 직무에서 검사도구의 타당도와 신뢰도를 실증적으로 재검토하여 보다 측정도구를 개선하고자 하는 노력이 필요할 것이다. 한편, 인적자원개발 분야 내 타당화 연구를 위한 제언은 다음과 같다. 첫째, 요인분석 시 주성분 분석과 직각회전의 사용은 지양해야 한다. 주성분 분석은 많은 변수를 소수의 주성분으로 축소하는 기법으로 요인분석과 그 목적이 다르고, 직각회전은 요인 간 상관을 0으로 가정하는데, 사회과학에서 변수 간 상관성이 0인 경우는 사실상 존재하지 않기 때문에 적합한 방법이 아니다. 따라서 주축분해법(principal axis factoring) 또는 최대우도법(maximum likelihood) 추정 방법과 사각회전(direct oblimin, promax, geomin 등)을 사용하고, 요인 수의 결정을 위해 평행선 분석(parallel analysis)을 고려하는 것을 권장한다. 특히 최대우도법과 완전정보 최대우도법은 결측치가 있는 자료를 처리하는 데 강점이 있다(Newman, 2003). 둘째, 인적자원개발 분야 내 보다 객관적인 검사도구 개발을 위해 후속연구에서 Rasch 모형의 활용을 권장한다. 본 연구에서 보여주었듯이, Rasch 모형을 활용하면 문항의 적합도, 곤란도, 문항 및 응답자 신뢰도, 응답 범주의 적절성 등을 추가적으로 검토할 수 있다. 기존의 요인분석이나 신뢰도 분석으로는 문제가 없는 문항이더라도 문항의 적합도, 곤란도, 응답 범주의 적절성에 문제가 있을 수 있음을 본 연구는 실증적으로 보여주었다.

마지막으로 본 연구는 보험설계사를 대상으로 수집된 이운수, 조대연, 김재민 (2018)의 자료를 활용하였다는 점에서 연구 결과를 일반화하는 데 한계가 있다. 또한 선행연구 검토에서 발견되었듯이, 조대연, 박용호(2011)의 측정도구에 대한 요인분석 결과에서 문제가 있었던 문항들은 연구들마다 상이했다는 점에서 본 연구에서 문제가 있다고 제기한 2문항 이외에도 다양한 관점에서 문항들이 재검토될 필요성이 있다. 끝으로 본 연구에서 사용한 자료에서 대부분이 긍정적으로 응답했다는 점은 자신의 관리자에 대한 평가에 있어 사회적으로 바람직한 편향적 응답이 발생했을 가능성을 배제할 수 없기 때문에 이를 고려한 후속연구가 축적될 필요가 있다.

참고문헌

- 강소운(2013). A항공사 리더의 코칭행동이 신입승무원의 직무몰입에 미치는 영향. 석사학위논문, 고려대학교 교육대학원.
- 강영순, 김정훈(2013). 코칭리더십, 기본심리욕구, 혁신행동, 조직시민행동의 구조적 관계. 대한경영학회지, 26(7), 1909-1928.
- 김선우(2018). 유통회사 구성원이 인식한 관리자 코칭행동과 구성원의 혁신행동의 관계. 석사학위논문, 고려대학교 교육대학원.
- 김세은(2018). S사 구성원의 경력계획 실천준비행동과 주관적 경력성공의 관계에서 관리자 코칭행동의 조절효과. 석사학위논문, 고려대학교 교육대학원.
- 김재민(2017). D손해보험 설계사가 인식하는 관리자 코칭행동과 주관적 경력성공의 관계. 석사학위논문, 고려대학교 교육대학원.
- 김지현(2012). 제조업 S기업 관리자의 코칭행동과 부하직원의 조직몰입 관계에서 경력계획 실천의 매개효과. 석사학위논문, 고려대학교 교육대학원.
- 김철영(2015). 국내 기업구성원의 조직정치지각과 조직몰입과의 관계에서 관리자 코칭행동의 조절효과. 석사학위논문, 고려대학교 교육대학원.
- 김현근(2016). S사의 관리자 코칭행동과 구성원 열의의 관계에서 구성원이 인식한 조직문화 유형별 수준의 조절효과. 석사학위논문, 고려대학교 교육대학원.
- 김홍준(2015). 출판유통업 관리자의 코칭행동과 조직구성원의 고객지향성의 관계에서 자기효능감의 매개효과. 석사학위논문, 고려대학교 교육대학원.
- 남기운(2017). 인수합병기업 관리자의 코칭행동과 구성원 직무성과와의 관계. 석사학위논문, 고려대학교 교육대학원.
- 문수미(2017). 기업체 관리자 코칭행동과 구성원 혁신행동 관계에서 긍정심리자본의 매개효과. 석사학위논문, 고려대학교 교육대학원.
- 박근수, 김주후(2011). 코칭의 평생학습적 의미와 코칭스킬 구성요인 탐색, 한국지방자치학회보, 23(4), 231-255.
- 박은영, 박소연(2008). 라쉬 모형 적용을 통한 양육태도 검사 양호도 분석, 지체중복건장장애연구, 51(4), 117-136.
- 박주훈(2012). H그룹 구성원의 핵심가치 전개활동과 조직몰입의 관계에서 관리자 코칭행동의 조절효과. 석사학위논문, 고려대학교 교육대학원.
- 성세실리아, 조대연(2016). HRD 분야에서 코칭 관련 연구 동향 탐색. 역량개발학습연구, 11(3), 1-24.
- 양유정(2016). L기업 관리자 코칭행동과 구성원 조직몰입 관계에서 조절조점의 매개효과. 석사학위논문, 고려대학교 교육대학원.

- 오효성(2009). 관리자의 코칭행동이 직원의 직무 만족에 미치는 영향에 대한 연구: 조직몰입과 자아 효능감의 매개효과를 중심으로. 석사학위논문, 숙명여자대학교 사회교육대학원.
- 이근원(2014). 중소기업 관리자 코칭행동이 구성원의 자기효능감에 미치는 영향. 석사학위논문, 고려대학교 교육대학원.
- 이운수, 조대연, 김재민(2018). 관리자 코칭 행동이 주관적 경력 성공에 미치는 영향: 횡단 다층 모형의 적용. HRD연구, 20(1), 225-247.
- 이운수, 정희원, 조대연(2016). 리더-구성원 간 공유 리더십이 팀 몰입 및 팀 만족에 미치는 영향: 행위자-상대자 상호의존 모형(APIM)의 적용. 기업교육연구, 18(2), 113-130.
- 이익섭, 홍세희, 신은경(2007). 장애 정체감 척도의 Rasch 모형 적용. 한국사회복지학, 59(4), 273-296.
- 이지아(2015). 사회초년생의 평가염려 완벽주의 성향과 인지적 유연성이 직무만족에 미치는 영향: 상사의 코칭행동을 매개로. 석사학위논문, 한국상담대학원대학교.
- 이지혜, 오상진(2017). 기업에서의 관리자 코칭 행동이 조직 구성원의 경력계획 실천행동에 미치는 영향. 경영교육연구, 32(2), 315-337.
- 이한중(2016). H손해보험 관리자 코칭행동과 구성원 혁신행동의 관계에서 심리적 임파워먼트의 매개효과. 석사학위논문, 고려대학교 교육대학원.
- 이혜정(2014). The mediation effect of self-directedness based on the relationship between managerial coaching behaviors and employee's job performance. 석사학위논문, 고려대학교 교육대학원.
- 이희권(2014). 직무특성과 직무소진 관계에서 관리자 코칭행동의 조절효과. 석사학위논문, 고려대학교 교육대학원.
- 정용성(2018). 중국 현채인의 관리자 국적에 따른 관리자 코칭행동과 심리적 안전감의 관계. 석사학위논문, 고려대학교 교육대학원.
- 전영화(2016). 한국과 중국의 전자 대기업에서 관리자의 코칭행동과 부하의 창의적 행동 관계 비교. 석사학위논문, 고려대학교 교육대학원.
- 조대연, 김희영(2009). 비즈니스 코칭 프로세스와 코칭 행동 탐색. 평생교육·HRD연구, 5(4), 51-71.
- 조대연, 박용호(2011). 관리자의 코칭행동과 부하직원이 인식한 직무성과의 관계. HRD연구, 13(4), 89-109.
- 조성진, 정이수(2018). 국내 코칭(Coaching) 연구 동향 및 향후 연구 방향. HRD연구, 21(3), 249-313.
- 조은현(2011). 코칭리더십 척도 개발 및 타당화와 코칭리더십이 조직태도에 미치는 영향. 박사학위논문, 광운대학교 대학원.
- 주두리(2016). C기업의 관리자 코칭행동과 조직몰입의 관계에서 팀 응집력의 매개효과. 석사학위논문, 고려대학교 교육대학원.

- 홍세희, 이현정, 손수경, 김효진, 윤미리, 강운경, 조기현(2018). 청소년 회복탄력성 척도의 Rasch 모형 분석. *미래청소년학회지*, 15(2), 107-124.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York, NY: Routledge.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation model* (pp. 136-162). Newbury Park, CA: Sage.
- DeMars, C. (2010). *Item response theory*. Oxford, England: Oxford University Press.
- Ellinger, A. D., Ellinger, A. E., & Keller, S. B. (2003). Supervisory coaching behavior, employee satisfaction, and warehouse employee performance: A dyadic perspective in the distribution industry. *Human Resource Development Quarterly*, 14(4), 435-458.
- Ellinger, A. E., Ellinger, A. D., & Keller, S. B. (2005). Supervisory coaching in a logistics context. *International Journal of Physical Distribution & Logistics Management*, 35(9), 620-636.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, 4(3), 272.
- Hong, S., Kim, B. S., & Wolfe, M. M. (2005). A psychometric revision of the European American Values Scale for Asian Americans using the Rasch model. *Measurement and Evaluation in Counseling and Development*, 37(4), 194-207.
- Hong, S., & Wong, E. C. (2005). Rasch rating scale modeling of the Korean version of the beck depression inventory. *Educational and psychological measurement*, 65(1), 124-139.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- Rasch, G. (1960). *Studies in mathematical psychology: Probabilistic models for some intelligence and attainment tests*. Oxford, England: Nielsen & Lydiche.

논문접수 2019. 06. 30 / 수정 2019. 08. 10 / 게재확정 2019. 08. 15

- 이윤수: 고려대학교에서 학사와 석사를 취득하고 미국 펜실베이니아 주립대학교에서 HRD

전공으로 박사과정을 수료하였음.

- 조대연: 미국 오하이오 주립대학교에서 박사학위를 취득하고, 현재 고려대학교 교육학과 부교수로 재직 중임.

Abstract

A Re-examination of Supervisor's Coaching Behavior Measurement using Rasch Rating Scale Model

Lee, Yunsoo (Korea University)

Cho, Daeyeon (Korea University)

The purpose of this study is to provide suggestions to improve supervisor's coaching behavior measurement developed by Cho and Park (2011) more objectively. To achieve this goal, we examined in various ways: not only factor analysis but also MNSQ (i.e., infit and outfit), item difficulty, and item category of the measurement using the Rasch rating scale model based on the Item Response Theory (IRT). The results showed that 1) the item difficulty of the current measurement needs to be revised, 2) the theoretical and practical discussion on whether the two-factor structure is appropriate is required, 3) it is required to improve the items that their infit and outfit seem inappropriate, and 4) it is necessary to make a reasonable decision about the item category of the measurement. Based on the results, we discussed theoretical and practical implications in the field of human resource development, provided suggestions for future studies, and emphasized the need for using the Rasch model.

[Keywords] Supervisor's coaching behavior, rasch model, rating scale model