

確認的因子分析による360度サーベイの特徴の把握

～自己評定と上司・同僚・部下による他者評定の測定等価性の検証～

前田純子 入江崇介

(株式会社 リクルートマネジメントソリューションズ 組織行動研究所 アセスメント開発部)

【背景と目的】

企業で人材育成や能力開発のために用いられる手法に360度サーベイがある。360度サーベイでは、対象者本人による自己評定と上司・同僚・部下というさまざまな仕事上の関係者からの他者評定が行われる。能力開発の研修では、自己評定と他者評定との結果の差異、さらには上司・同僚・部下という評定者グループによる評定結果の差異に着目し、個人の強み・弱みや周囲の期待を確認するきっかけとすることが多い。

このようなさまざまな評定者グループによる評定結果の単純比較が対象者本人に有益な気づきを促すためには、異なる背景をもった評定者グループによって行われるサーベイの結果について測定等価性が成り立つことが前提となる。すなわち、特定の評定者グループにおいて、評定結果の高低差や仮定される尺度間の相関関係などに特異な傾向がないことがこの前提に当たる。

この360度サーベイの測定等価性について検証を行った研究の代表的なものはDiefendorff et al.(2005)である。この研究ではVandenberg & Lance(2000)で提案された確認的因子分析(CFA)の手法を活用し、360度サーベイの概念的(conceptual)等価性と心理測定的(psychometric)等価性の詳細な検証を行っている。しかしながら、日本においては360度サーベイの測定等価性についてこのような詳細な検討は行われてこなかった。

本研究では、日本のマネジメント層を対象にした360度サーベイについて、本人と上司・同僚・部下という4つの異なる評定者グループのデータを用いてDiefendorff et al.(2005)の方法を参考にその測定等価性の検証を行うことを目的とする。

【方法】

使用尺度：複数観察者評価システムM型(Multi-Observer Assessment M型; MOA-M 株式会社リクルートマネジメントソリューションズ)を用いた。この360度サーベイは、課題形成力、課題遂行力、人材活用力、対人対応力の4尺度から成り、1尺度あたり16項目の計64項目で構成されている。評定形式は5段階の両側項目である。

分析対象：38社24979名分のデータのうち、自己・上司・同僚・部下のそれぞれのグループにおいて、64項目全てに回答している評定者が1名以上得られているデータを分析対象として抽出した。その結果、37社9436名分(対象者2359名分)のデータが得られた。なお、上司・同僚・部下の評定者グループにおいてそれぞれ2名以上の評定者が存在する場合は、ランダムに1名を選択した。

手続き：本研究においては、Diefendorff et al.(2005)と同様に、本人・上司・同僚・部下という異なる評定者グループの評定結果を複数の母集団ではなく単一の母集団から得られるデータとして分析

を行った．これは，それぞれの評定者グループに属する評定者が単一の対象者について評定を行うという360度サーベイの特徴を反映したものである．

なお，分析にはAMOS 4.0 Graphicsを用いた．その際，ソフトウェアの制約上，64項目×4評定者グループ=256項目のデータを分析できないことを考慮し，本研究においては1尺度あたり16項目から8項目に削減したデータを使って分析を行った．分析に用いる項目の抽出に当たっては，自己評定のデータを用いて斜交プロクラステス解による因子分析を行い，その結果それぞれの尺度に対して負荷量が大きな項目を8つずつ選択するという方法をとった．その結果得られた尺度の係数は課題形成力0.870，課題遂行力0.871，人材活用力0.867，対人対応力0.827であった．

検証するモデルの説明：Vandenberg & Lance(2000)に提案されたCFAによる手法に従い，以下の8つのモデルの検証を段階的に行った．その詳細は表1に示す．

表1：本研究で検証を行った各モデルの等置条件

モデル	モデルの概要 (そのモデルで追加される制約)	グループ間の等置条件	帰無仮説が 棄却された場合の示唆
モデル0	観測変数間の分散・共分散が 全てのグループで等しい	観測変数間の分散・共分散	いずれかのグループで 測定等価性が成り立たない
モデル1	各因子を定義する観測変数が 全てのグループで等しい	観測変数と因子の対応	いずれかのグループで 因子の数や構造に違いがある
モデル2	各因子の因子パタンが 全てのグループで等しい	観測変数と因子の対応，因子負荷量	いずれかのグループで 因子を代表する項目などに違いがある
モデル3	観測変数の切片が 全てのグループで等しい	観測変数と因子の対応，因子負荷量 切片	いずれかのグループで 評定得点に違いがある
モデル4	誤差分散が 全てのグループで等しい	観測変数と因子の対応，因子負荷量 切片，誤差分散	いずれかのグループで 尺度の信頼性が異なる
モデル5	因子の分散が 全てのグループで等しい	観測変数と因子の対応，因子負荷量 切片，誤差分散，因子分散	いずれかのグループで 尺度の評定レンジが異なる
モデル6	因子間の共分散が 全てのグループで等しい	観測変数と因子の対応，因子負荷量 切片，誤差分散，因子分散 因子間共分散	いずれかのグループで 因子間の関連に違いがある
モデル7	因子得点が 全てのグループで等しい	観測変数と因子の対応，因子負荷量 切片，誤差分散，因子分散 因子間共分散，因子得点	いずれかのグループで 潜在構造で評定傾向に違いがある

この方法では，段階的に評定者グループ間に，より厳しい等置制約を仮定することによって，詳細に測定等価性の検証を行うことができる．なお，それぞれの制約を置いた際に，あてはまりが極端に悪くなる場合の影響要因を確認するために，下位テストとして尺度ごとや評定者グループごとの分析を行うことが有効である．本研究では，総合的にいずれかの評定者グループに特異な傾向があるかを確認するために，1つの評定者グループのみ特異な傾向を有するというモデルの検証もあわせて行った．

【結果】

表2に，モデル0~7について得られた各種適合度指標を示す．また，表3にモデル2~7において，1つの評定者グループについて相等性の制約を外したモデルとの自由度と2乗値の変化量を示す．

表2: モデル0~7についての適合度指標

モデル	モデルの比較	df	2乗値	RMSEA	TLI	SRMR	CFI	df	2乗値	CFI
モデル0		7728	16381.905	0.022	0.952	0.103	0.954			
モデル1		7880	22411.874	0.028	0.921	0.038	0.923			
モデル2	モデル2 - モデル1	7964	22771.389	0.028	0.920	0.038	0.922	84	359.515	-0.001
モデル3	モデル3 - モデル2	8048	24089.052	0.029	0.915	0.038	0.915	84	1317.663	-0.007
モデル4	モデル4 - モデル3	8144	25803.472	0.030	0.907	0.039	0.907	96	1714.420	-0.008
モデル5	モデル5 - モデル4	8156	26374.607	0.031	0.904	0.053	0.904	12	571.135	-0.003
モデル6	モデル6 - モデル5	8174	26488.004	0.031	0.904	0.053	0.903	18	113.397	-0.001
モデル7	モデル7 - モデル6	8186	27182.332	0.031	0.900	0.053	0.900	12	694.328	-0.003

2乗値の変化量はいずれも $p < 0.01$ で有意.

表3: 1つの評定者グループの相等性の制約を外した場合の 2乗値の変化

モデル	df	相等性の制約を外した評定者グループ			
		本人	上司	同僚	部下
		2乗値	2乗値	2乗値	2乗値
モデル2	28	229.043	77.434	54.591	99.060
モデル3	28	920.124	301.513	152.067	355.362
モデル4	32	380.205	574.843	310.753	1042.503
モデル5	4	419.850	26.681	26.657	270.398
モデル6	6	34.456	27.697	33.232	58.204
モデル7	4	356.362	168.745	108.580	212.976

2乗値の変化量はいずれも $p < 0.01$ で有意.

結果1: モデル全体のあてはまり (表2参照)

モデル0の適合度はVandenberg and Lance(2000)で採用されている適合度の基準(RMSEA 0.08, TLI 0.90, SRMR 0.15, CFI 0.90)を満たした。その場合、モデル1~7も成り立つことが予想されるため、それ以降のモデルの検証は必要ないとされている。しかし、Diefendorff et al.(2005)同様、モデル1~7についての確認を行うことによってCFIの特性を活かして詳細に評定者グループの特異性について検証を行うことができるため、モデル1~7についても検証を行った。

その結果、モデル1~7のいずれにおいても、各種適合度指標は推奨される基準を満たした。なお、1つ前のモデルとの2乗値の変化量は有意であった(いずれも $p < 0.01$)が、CFIの変化量は0.01未満であることから、より制約の強いモデルについてもそれが棄却されないことを意味する(Cheung and Rensvold, 1999)。よって、最も制約の強いモデル7が成り立ち、MOAMにおいて本人・上司・同僚・部下という4つの異なる評定者グループ間で測定等価性が成り立つことが示された。

結果2: 個別の評定者グループの相等性の制約を外した場合のモデル適合度の変化(表3参照)

2乗値の変化量は, モデル2では「本人」, モデル3では「本人」, モデル4では「部下」, モデル5では「本人」, モデル6では「部下」, モデル7では「本人」について相等性の制約を外した際に最も大きくなった(いずれも, $p < 0.01$ で有意). すなわち, それぞれ表1で説明された等置条件についての特異性がこれらの評定者グループにあることが示唆される.

【考察】

本研究では, Diefendorff et al.(2005)同様に, 自己評定と上司・同僚・部下の他者評定においてツールの測定等価性の仮定を置くことが可能であることが示された. すなわち, 自己評定と上司・同僚・部下において結果に差があった場合, それは評定構造の違いなどの認知の違いによるものではなく, 対象者に対する評定者の観察結果の差として解釈することが可能であることが改めて確認された. このことから, 自己評定を含むさまざまな評定者グループの評定結果を比較することは能力開発の研修などで自己の気づきを促す上で有用なことであり得ると考えられる.

ただし, 等置制約を強めることによって有意な2乗値の変化があること, いずれかのグループのみのパラメータを自由に推定した方がモデルのあてはまりがよくなることから, 若干の評定傾向の差がある可能性が同時に示された. このことは, CFAの手法を用いて360度サーベイの測定等価性を検証することによって, 個別の項目や特定の評定者グループに起因する等価性の阻害要因が確認可能であることを示している. また, 具体的な検討を行うためには, 平均値や相関係数などを確認することも必要である.

しかし, 単純な統計的有意性の確認だけでは本質的な測定の等価性・特異性について結論を出すことは難しい. したがって, 今後は本研究の成果を踏まえて, 各評定者グループに特有だと考えられる評定特性を考慮しながら, 評定者グループごとの測定の等価性と特異性について具体的にその確認を行っていくことが必要である. それによって, 360度サーベイの性質についてより深い理解が可能になり, その実用のために有効な知見が得られると考えられる.

【参考文献】

Cheung, G.W., & Rensvold, R.B. 1999 What constitutes significant differences in evaluating measurement invariance? Paper presented at the 1999 conference of the Academy of Management, Chicago.

Diefendorff, J.M., Silverman, S.B., & Greguras, G.J. 2005 Measurement equivalence and multisource ratings for non-managerial positions: Recommendations for research and practice. *Journal of Business and Psychology*, 19, 399-425.

Vandenberg, R. J., & Lance, C. E. 2000 A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-69.