

多面観察評価における 上司・同僚・部下の評定結果の統合について ～多変量一般化可能性理論による一般化可能性の検証～

入江崇介 鷲坂由紀子 舛田博之 二村英幸

(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)

【背景と目的】

経営人事の現場において人材の育成ならびに評価を目的として用いられる手法に多面観察評価がある。多面観察評価では、さまざまな立場の複数の関係者からの他者評定が、育成の観点では本人の「気づき」を促すために、評価の観点では情報の客観性と信頼性を増すために、重要な役割を果たす。この関係者は、評定の対象者とは上司・同僚・部下などさまざまな関係にあるため、たとえば階層毎に評定の平均値をとり、それが研修の際のフィードバックや評価の参考情報に用いられることがある。このような階層などのグループ毎の平均値の一般化可能性については先行研究で検討がなされている (e.g. Greguras et al., 1998; 入江ら, 2004)。

しかし実際には、それぞれのグループに属する評定者数が十分ではない場合や、さまざまなグループの関係者からの評定結果の総体に注目する場合に、全関係者の評定結果を統合した重み付き和得点が用いられることもある。このような多様なグループに属する複数の関係者の評定を統合した重み付き和得点の一般化可能性を検証した研究はこれまで行なわれていない。

この複数の特性や階層のようなカテゴリーに属する変数の重み付き和得点の一般化可能性を推定する方法に多変量一般化可能性理論

(Multivariate Generalizability Theory ; Cronbach et al., 1972) がある。そこで本研究では、多変量一般化可能性理論を用いて、対象者と評定者との関係の多様性を考慮に入れた評定の一般化可能性を検証した。

【方法】

手続き: 大部分のサンプルにおいて対象者毎に評定者が異なるため、 $r:t$ デザインに基づき分散・共分散成分の推定を行なった。実験計画の概念図は Figure 1 である。

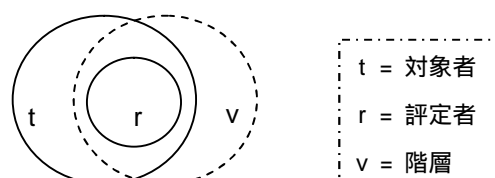


Figure 1 $r:t$ デザインの概念図

このモデルは、対象者毎に異なる評定者の相が対象者の相にネストし、評定者の相はそれぞれの人がある階層の相にネストし、対象者はあらゆる階層の評定者から評定を得るために対象者の相と階層の相はクロスしていることを示している。なお、評定者が属する階層は上司・同僚・部下の3つの水準に固定されているため、母数として扱われる。分析には mGENOVA (Brennan, 2001) を用いた。

使用尺度：複数観察者評価システム MOA-M (Multi-Observer Assessment M型, (株)リクルートマネジメントソリューションズ)を用いた。このツールは、課題形成、課題遂行、人材活用、対人対応の4尺度からなり、1尺度あたりの項目数は16である。各尺度の係数は、課題形成 = 0.951、課題遂行 = 0.950、人材活用 = 0.947、対人対応 = 0.940と極めて高い水準であった。また、MOA-Mにおいて各尺度の項目数が16の場合は項目内容の差異に起因する変動の影響は極めて小さいことが確認されている(入江ら, 2004)。よって、今回は各尺度について16項目の平均点である尺度得点を算出し、それをもとに分析を行った。それにより、項目の相をモデルから除外することができ、mGENOVAでの分析が可能になった。

分析対象：上司・同僚・部下のそれぞれの階層において、64項目全てに回答している評定者が2名以上得られているデータを分析対象とした。それぞれのグループの評定者が3名以上存在する場合は、ランダムに2名を抽出した。その結果、38社24979名分の評定者データから、28社4062名分のデータ(対象者677名分のデータに相当)が抽出された。

【結果】

(1) 多変量一般化可能性研究の結果

4つの尺度のそれぞれについて、分散・共分散成分の推定を行った。その結果を Table 2 - Table 5 に示す。

Table 2 「課題形成」における
分散・共分散成分の推定結果

$$\hat{\Sigma}_t = \begin{bmatrix} 0.270 & & \\ 0.178 & 0.132 & \\ 0.175 & 0.163 & 0.239 \end{bmatrix}$$

$$\hat{\Sigma}_{r:t} = \begin{bmatrix} 0.313 & & \\ & 0.379 & \\ & & 0.388 \end{bmatrix}$$

Table 3 「課題遂行」における
分散・共分散成分の推定結果

$$\hat{\Sigma}_t = \begin{bmatrix} 0.257 & & \\ 0.174 & 0.149 & \\ 0.168 & 0.177 & 0.274 \end{bmatrix}$$

$$\hat{\Sigma}_{r:t} = \begin{bmatrix} 0.293 & & \\ & 0.370 & \\ & & 0.391 \end{bmatrix}$$

Table 4 「人材活用」における
分散・共分散成分の推定結果

$$\hat{\Sigma}_t = \begin{bmatrix} 0.214 & & \\ 0.139 & 0.115 & \\ 0.130 & 0.143 & 0.219 \end{bmatrix}$$

$$\hat{\Sigma}_{r:t} = \begin{bmatrix} 0.301 & & \\ & 0.389 & \\ & & 0.453 \end{bmatrix}$$

Table 5 「対人対応」における
分散・共分散成分の推定結果

$$\hat{\Sigma}_t = \begin{bmatrix} 0.160 & & \\ 0.123 & 0.116 & \\ 0.129 & 0.161 & 0.237 \end{bmatrix}$$

$$\hat{\Sigma}_{r:t} = \begin{bmatrix} 0.304 & & \\ & 0.374 & \\ & & 0.405 \end{bmatrix}$$

ただし

$$\hat{\Sigma} = \begin{bmatrix} 2 & & & \\ \text{上司} & & & \\ \text{上司} \cdot \text{同僚} & 2 & & \\ \text{同僚} & & & \\ \text{上司} \cdot \text{部下} & \text{同僚} \cdot \text{部下} & 2 & \\ \text{部下} & & & \end{bmatrix}$$

であり、行列の対角要素はそれぞれの階層における分散成分、非対角要素は階層間の共分散成分である。また、 $\hat{\mu}_i$ は対象者の主効果、 $\hat{\mu}_{r:t}$ は評定者の主効果・対象者と評定者の交互作用・その他の誤差に起因する成分である。

(2) 多変量決定研究の結果

4つの尺度のそれぞれについて、実際の実施状況にあわせて各階層の評定者数に応じて重み付けを行なった和得点(e.g. 上司1名、同僚1名、部下2名の場合の重み; 上司=0.25、同僚=0.25、部下=0.5、ただし上司+同僚+部下=1とする)と、実施の際の評定者数よりも各階層からの評定を等しく扱うことを重視して1/3ずつ重み付けを行なった和得点の多変量一般化可能性係数(Multivariate Generalizability Coefficient、以下MGC)を算出した。なお、MGCは、以下の式(1)によって求められる。

式(1)

$$MGC^2 = \frac{\sum_t \sum_{t'} w_t w_{t'} \sigma_{tt'}^2}{\sum_t \sum_{t'} w_t w_{t'} \sigma_{tt'}^2 + \sum_t w_t^2 \sigma_{r:t}^2 / n_{r:t}}$$

wは重み、 $\sigma_{tt'}$ は分散・共分散成分、nは評定者数を表す。

すなわち、MGCは複合された母得点の分散と、複合された観測値の分散の期待値との比によって定義され、信頼性係数に準じた解釈が可能である。また、それぞれの階層における評定者数 $n_{r:t}$

を変化させることによって、さまざまな実施状況下でのMGCをシミュレーションすることが可能となり、これを多変量決定研究という。4つの尺度毎にとについて多変量決定研究を行なった結果をTable 6 - Table 9に示す。

Table 6 「課題形成」における評定者数とMGC

評定者数			多変量一般化可能性係数	
上司	同僚	部下	評定者数に応じた重み付け	1/3ずつ重み付け
1	1	1	0.608	0.608
1	1	2	0.676	0.654
1	2	1	0.653	0.652
1	4	3	0.785	0.757
2	1	1	0.696	0.644
2	2	2	0.756	0.756
2	2	3	0.784	0.779
3	3	3	0.823	0.823

Table 7 「課題遂行」における評定者数とMGC

評定者数			多変量一般化可能性係数	
上司	同僚	部下	評定者数に応じた重み付け	1/3ずつ重み付け
1	1	1	0.620	0.620
1	1	2	0.691	0.667
1	2	1	0.668	0.664
1	4	3	0.800	0.769
2	1	1	0.702	0.654
2	2	2	0.765	0.765
2	2	3	0.794	0.788
3	3	3	0.830	0.830

Table 8 「人材活用」における評定者数とMGC

評定者数			多変量一般化可能性係数	
上司	同僚	部下	評定者数に応じた重み付け	1/3ずつ重み付け
1	1	1	0.545	0.545
1	1	2	0.617	0.599
1	2	1	0.598	0.591
1	4	3	0.744	0.714
2	1	1	0.638	0.580
2	2	2	0.706	0.706
2	2	3	0.737	0.734
3	3	3	0.783	0.783

Table 9 「対人対応」における評定者数とMGC

評定者数			多変量一般化可能性係数	
上司	同僚	部下	評定者数に応じた重み付け	1/3ずつ重み付け
1	1	1	0.553	0.553
1	1	2	0.639	0.603
1	2	1	0.608	0.599
1	4	3	0.764	0.716
2	1	1	0.626	0.590
2	2	2	0.712	0.712
2	2	3	0.750	0.739
3	3	3	0.788	0.788

なお、本研究で用いた38社24979名分の評定者データにおいては(a)上司2名、同僚2名、部下3名、Greguras(1998)においては(b)上司1名、同僚4名、部下3名という評定者数が一般的であった。

MGCは(a)の場合、評定者数に応じた重み付けを行った際、課題形成=0.784、課題遂行=0.794、人材活用=0.737、対人対応=0.750であった。また、1/3ずつ重み付けを行なった際は課題形成=0.779、課題遂行=0.788、人材活用=0.734、対人対応=0.739であった。

(b)の場合、評定者数に応じた重み付けを行った際は課題形成=0.785、課題遂行=0.800、人材活用=0.744、対人対応=0.764であった。また、1/3ずつ重み付けを行なった際は課題形成=0.757、課題遂行=0.769、人材活用=0.714、対人対応=0.716であった。

いずれの場合も、MGCは0.7を超えていた。ただし、全般的に課題に関わる側面に比べ対人関係に関わる側面においてややMGCが低くなる傾向にあった。また、ほとんどの場合、評定者数に応じた重み付けを行った場合に比べ、1/3ずつ重み付けを行なった場合にMGCが低くなる傾向にあった。

【考察】

本研究では、上司・同僚・部下という3つの階層の評定者による評定結果を統合した際の一般化可能性についての確認を行なった。本抄録掲載以外の評定者数のパターンについても確認したところ、上司・同僚・部下の合計が7名程度であれば、評定者数に応じた重み付けを行なった重み付き和得点には0.7程度の十分な一般化可能性があることが確認された。ただし、評定者数によっては1/3ずつ重み付けを行なった際にはやや一般化可能性が低く0.6程度になることが確認された。この傾向はある階層の評定者が1名のよ

うに極端に少ない場合に顕著であった。

また、階層毎の分散成分ならびに階層間の共分散構造が異なるため、評定者の合計人数が同じでもMGCが異なった。このことも、人数に応じた重み付けを行なうか、等しく重み付けを行なうかと同様に最終的な結果の一般化可能性に影響を与える一因であった。

以上より、多面観察評価における評定結果の統合において、単純に全体の総和を求める場合、すなわち階層毎に評定者数に応じた重み付けをして和得点を求める場合と、それぞれの階層毎の評定を等価に扱うために等しい重みをつけて和得点を求めるなど任意の重み付けを行なって和得点を求める場合とで、その一般化可能性が異なることが確認された。階層毎に集計した結果を等しい重みで統合した際には、極端に評定者数が少ない階層があった場合に、その得点の信頼性がやや低くなることに留意する必要がある。

【参考文献】

Brennan, R.L. (2001). *Manual for mGENOVA*. Version 2.01. Iowa: Iowa Testing Programs, University of Iowa.

Cronbach, L.J., Glesser, G.C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles*. New York: Wiley.

Greguras, G.J., & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. *Journal of Applied Psychology*, 83, 960-968.

入江崇介, 鷺坂由紀子, 舛田博之, 二村英幸 (2004). 多面観察評価手法の特性に関する検討～一般化可能性理論によるアプローチ～. 日本テスト学会第2回大会発表論文抄録集, 68-69.