

モニタリングの負担と人事評価の歪み

—自己評価・上司評価の相関—

The Distortion of the Personnel Evaluation Caused by the Difficulty of the Monitoring

— The Correlation of the Rating by the Subordinates and Their Bosses —

中嶋哲夫*

Tetsuo NAKASHIMA *

Abstract

The purpose of this paper is to describe the distortion that occurs in the personnel evaluation by the difficulty of the monitoring.

To realize this purpose, we use the data set of 1 enterprise which are comprised the interview investigation data, a questionnaire survey and the personnel evaluation data.

The results are the following; 1) when the monitoring is difficult, the correlation of the rating which a subordinate evaluated, and the rating which his boss evaluated, becomes small, 2) the rating which his boss evaluated tends to be distorted toward leniency or central tendency.

キーワード：自己評価、上司評価、寛大化、中心化、モニタリング負担

Keywords : evaluation by subordinates, evaluation by their bosses, leniency, central tendency, difficulty of the monitoring

* 人事教育コンサルタント (大手前大学非常勤講師)

1. 問題の所在

1990年代後半以降、日本企業は「成果主義」への人事制度改革を精力的に進めている。その内容にはさまざまな形態があるが、一連の改革に共通するのは、1) 目標の達成度やコンピテンシーなどを評価基準にして、短期的なパフォーマンスを評価し、2) 面接制度を通じて、査定結果に対する従業員の納得性を確保したうえで、3) 賃金の格差を広げる、という方向であろう。ここでは、面接制度による上司と部下の合意形成によって、評価の納得性を高めようとする意図に注目しておきたい。

しかし、それらの取り組みが効果をあげているかといえ、必ずしもそうではないようである。例えば、平成16年度労働経済の分析によれば、納得性を確保した評価制度の導入に取り組んでいる企業においても60%強の人はそれが従前と変わらないとし、20%ほどの人はむしろ納得性が下がったと評価している。また、企業の側からも人事評価制度のもつさまざまな問題点が指摘されている(例えば雇用管理調査(2002))。城(2004)や柳下(2003)など実務側から成果主義的な人事評価の難しさを指摘する声も強い。総じて、オープンな人事評価制度を導入したものの、従業員の人事評価への納得度は高まっていないように思える。

このようなことが起きる原因のひとつとして、さまざまな特性をもつ職務に対して単一の評価制度を適用するために、効果的な制度運用ができないという問題があるだろう。例えば、MilkovichとBoudereau(1997)は職務のルーチン性、環境の安定性、従業員の独立指向性の3つを軸とする空間に様々な評価方法を位置づけ、職務特性によって異なる人事評価がふさわしいことを主張している。同様に、MilgromとRoberts(1992)は、業績測定の難易度によってインセンティブの強度を異なったものをする必要があることを指摘している。また、国内においては奥野(2004)が、1企業における目標管理制度の運用状況を調査し、うまく運用できる職種とそうでない職種があることを明らかにし、それが人事評価制度の納得度とも関係していることを明らかにしている。加えて中嶋(2007)は人事評価制度の納得度がマネジメント手続きの実施度合いによって異なり、しかも、職種によって納得度に影響を与えるマネジメント手続きが異なることを明らかにしている。これらの研究は、人事評価の納得度が高まらない原因として、職務特性と評価制度のマッチングの問題があることを示唆している。企業内の職務がさまざまな特性を持っているとすれば、それを単一の人事評価制度によって評価しようとすることに無理があるのかもかもしれない。

いっぽう、人事マイクロ・データを用いて、評価者負担の観点から人事評価制度の運用を説明する研究も行われている(中嶋・梅崎・松繁(2004)、梅崎・中嶋(2005)、中嶋(2004)など)。彼らの研究によれば、評価者負担の大きさによって管理者が機会主義的な行動を行い、それによって人事部が意図しない査定結果や、制度運用が行われがちであることが明らかにされている。

以上から考えれば、単に面接制度を導入したからといって人事評価制度の納得度が高まるわけではないと考えることができる。評価制度の納得度を高めるためには、面接制度の運用やそれに関連

する評価者の負担にまで踏み込んだ研究をおこない、そこから実践的な施策を打ち出すことが必要であろう。そのためには、自己評価と上司評価がどの程度相関しているのか、どのような職務特性のもとで、上司と部下の認識ギャップが生まれやすいのか、それに対してどんな手を打つことができるのか、といった研究を行う必要があるだろう。

そこで本稿は、1) 自己評価と上司評価の一致度合いに着目し、2) それを生み出す職務特性（成果と努力過程の観察しやすさ）と、3) それが生み出す上司評価への影響（査定分布の傾向）を観察し、その因果関係を明らかにする。

本稿は次のように構成する。続く第2節では研究仮説を述べる。第3節では、対象企業とその人事評価制度の概要を説明する。第4節は、対象職種の職務特性の説明である。第5節では、自己評価と上司評価の相関係数を確認した後、上司評価の分布を把握し、その因果関係を考察する。第6節はまとめである。

2. 研究仮説

人事評価を行う管理者は、時間的な負担と心理的な負担のもとで評価行動を行っていると考えられる。その負担は評価者の機会主義的な行動を生み出し、結果として査定の分布が歪曲されたり、面接が行われなかったりすることにつながる可能性がある。例えば、中嶋・梅崎・松繁（2004）は、年齢給を廃止した企業において賃金が年功化した事例を発見し、その原因として制度改訂後に査定格差が縮小したことを明らかにしたうえで、評価者負担の増大がその原因であることを示唆した。また、梅崎・中嶋・松繁（2003）は、1次評価から最終評価までの評価の調整過程を分析し、そこにおいて評価の格差が変動することを明らかにし、部下に直接向かい合う評価者は厳しい評価を行いにくいことを示唆した。さらに、梅崎・中嶋（2005）は、人事部が定めた人事評価手続きが、どの程度まで実行されているのかを確認し、評価者負担が大きい場合には、それらが実施されにくくなることを明らかにした。また彼らは、人事評価手続きの実施ダミーを被説明変数としたプロビット推定を行い、上司への尊敬、上司と部下の年齢差の小ささ、評価結果の3つの説明変数が査定通知の実施にプラスの効果を与えることを明らかにした。このように、評価者負担によって、人事評価の分布が影響を受けたり、あるいは決められた手続きが実施されなかったりすることが明らかになりつつある。

いっぽう、中嶋（2004）は、1企業の営業所長について査定結果と生産性データを精査した。それにより、上司と異なるオフィスに勤務し、上司がモニタリングをしにくい地方の営業所長は、売上予算の規模や部下人数などの客観的な指標を根拠にして評価されがちであるのに対し、上司と同じオフィスに勤務する大都市の営業所長は、客観的な指標に依拠して評価される比率が小さいことを明らかにした。しかし、彼は、営業部長と同じ場所で勤務しているかどうかをモニタリングの容易さの代理指標としている。実際に、職務特性を明らかにしたとはいえないだろう。

上記の研究をふまえ、本稿では次の2段階の仮説を設定し、検証する。第1段階は、中嶋（2004）が指摘したモニタリングの容易さを、成果測定の容易さと努力観察の容易さに分割する。そのうえで、部下の自己評価と上司評価の一致度合いに各々が与える影響を分析する。もし成果分布の幅が大きく、成果が測りやすければ、業績評価において自己評価と上司評価は容易に一致するであろう。逆に、成果の分布幅が小さかったり、それが測りにくかったりすれば、自己評価と上司評価は、一致しにくいであろう。同様に、もし努力過程を観察しやすければ、行動評価において自己評価と上司評価の一致度合いは大きく、逆に、それを観察しにくければ、それは小さくなるであろう。

第2段階は、自己評価と上司評価の一致度合いが評価結果の分布に与える影響を分析する。もし、一致度合いが大きければ、上司は部下に査定を説明する負担を感じず、上司の査定は人事部が求める正規分布に近づく可能性が高い。いっぽう、一致度合いが小さければ、上司は部下に説明することを負担に思うだろう。この場合、査定の分布幅を縮小して評価の差を小さくし、部下への説明負担を軽減する行動を上司がとることも考えられる。また、同様に、上司が寛大な評価を行い、結果として分布が上方に偏ることも考えられる。このような場合、評価調整の最終段階で人事部門が介入し、査定を大きく変更することが考えられる。

以上の仮説を次のような手続きで立証する。まず、努力観察と成果測定の容易さを聞き取り調査とアンケート調査によって把握する。次に、自己評価と上司評価の一致度合いを確認するため、人事マイクロ・データからふたつの相関係数を計算する。第3に、上司評価の分布を人事マイクロ・データから把握する。最後に、それらの因果関係を考察する。

3. 対象企業の概要

3-1. 対象企業の概要と業務の特徴

E社はインテリア工事を業とする中規模企業である。バブル崩壊以降、業績低迷が続いていたが、近年、持ち直しつつある。1990年代の新規採用が少なく、平均年齢は47歳、平均勤続年数は26年となっている。学卒者の比率は30%、女性比率は5%である。従業員は400人弱である。

事業は完全受注型であるため業務内容にはいくつかの特徴がある。第1に、案件ごとに営業職、工務職、設計職など、複数職種のメンバーが相互に関わりながら仕事を進める。その関わりは、社員だけでなく外注業者にもおよぶ。第2に、受注案件の規模の大きさが極端に異なる。半日で終わるような店舗工事であれば、1年以上かかるビルの内装工事もある。その差が大きいため、担当する案件は日常的にきめ細かく調整される。第3に、建設工事の最終工程であるため、他工程の遅れがしわ寄せされる。いざとなれば、「仙台や博多から大工を呼び寄せてでも納期を守る」というアクションが必要となる。第4に、業務が機能別に組織されている。案件の規模や受注タイミングが多様であるため、機能別に組織編成し、案件ごとに担当者を決めて流動的なチームで仕事を進める方が効率的だからであろう。このため、管理者は「誰がどのような経験をしていて、いつ手が空い

ているのか」をしっかりと把握し、担当させる案件を調整する必要がある。したがって、管理者が部下の行動や成果を把握することが事業遂行上重要な業態であると考えることができよう。

3-2. 対象企業の人事評価制度

E社は2004年度に職能資格制度を改定し、従来秘密にされていた人事評価の基準や手続きを従業員に公開した。それによって、人事制度の透明度を増し、従業員の能力開発を促進するという目的が社員向けのパンフレットにうたわれている。

人事評価は3つの内容で構成されている。第1は業績評価、第2は行動評価、第3は能力評価である。業績評価と行動評価が目標管理と連動している。前者は目標の達成結果を、後者は目標への取り組み姿勢を評価する。ふたつの評価は1枚のシートにまとめられ、どちらも自己評価と直属上司による評価が書き込まれる。評価結果は昇給、賞与の決定に用いられる。能力評価は、教育、異動に活用される。こちらも自己評価と上司評価が行われる。昇格は3つの評価を勘案して実施される。

分析を行うのは業績評価と行動評価であるので、その手続きを詳述する。前述のように、ふたつの評価は1枚のシートを使って同時に行われる。期首には目標設定面接が、3ヶ月後には中間面接が、半期末には期末評価が行われる。

期末評価の手続きは次の通りである。まず自己評価が実施される。つぎに、部下の自己評価を参考にして上司が評価を行う。ふたつの上司評価はウェイトをかけて合算され、その結果が1次評価とされる。合算にあたっては上位等級では80%、下位等級は50%が業績評価とされ、残りの20%と50%が行動評価とされる。1次評価が行われた後、評価者が部門ごとに集まり、評価調整を行う。そこでは個別の評価内容を評価者が相互に吟味しあう。1次評価者の間で合意形成された結果を部門長が追認し、2次評価とする。最後に、部門長と人事部門が調整会議を開き、部門間の評価の甘辛を調整し、最終評価を確定させる。最終評価確定後に1次評価者が査定結果を部下に説明する¹⁾。

評価は、各々7段階で実施される。A+、A、B+、B、B-、C、C-の7段階である。Bが「求める基準にほぼ一致」とされている。分布規制などは行われていないが、貢献度に応じた差をつけることを人事部門が強調し、寛大な査定結果と見られた部門に対しては、最終段階で人事部門が直接介入し評価結果が調整されることもある。

評価者となるのは、当該組織の管理者である。平均して5.4人、最大は31人、最小は1人の部下を評価する。E社の組織は機能別に編成されているので、上司と部下の専門性は比較的一致している。また、「人事異動が少ないため、上司も部下も経験が長く、担当業務についてはお互いによく知っている」と人事担当者がいう。また、評価者に対しては、人事部が目標設定および評価に関する研

1) 人事担当者に確認したところ、人事部門が査定の説明を行わせるよう強力に指導をしたため、査定結果を部下に説明することは、ほぼ完全に実施されているとのことである。

修を開催し、制度の周知徹底を図っている。上司は仕事についても評価制度についても十分な知識を持っていると考えて良いだろう。

4. 職務特性の把握

本節では、聞き取り調査とそれをふまえたアンケート調査によって明らかになった職種ごとの職務特性を説明したい。ふたつの調査結果には若干の違いがあるが、それらに共通する結果を、より確からしい職務特性と捉えることにする。

4-1. 聞き取り調査による把握

最初に、2005年8月に行った聞き取り調査の結果を説明する²⁾。対象職種は統計的な検討が可能な6職種とした。すなわち、事務職、工務職、営業職、生産職、設計職、および部門管理者³⁾である。各々の職種について、管理者と一般従業員を1名ずつ人選し、各1時間程度の聞き取りを行った。対象者の人選は、1) 各職種に関して幅広く自職場以外の情報を持つ人物であること、2) 比較的偏らない意見を持つ人物であること、のふたつを基準としてE社人事部にお願いした。聞き取りには、次の4点に焦点を当てた半構造化インタビューを用いた。1) 成果の分布幅、すなわち、成果の分布が大きな幅をもっているのか、それとも狭い範囲に分布しているのか、2) 成果の測りやすさ、すなわち、何らかの客観的指標で成果を捉えることができるのか、それともそうでないのか、3) 努力観察の容易さ、すなわち上司が部下の行動を観察できる時間の長さ、4) 努力内容を評価ランクに当てはめる判断基準のわかりやすさ、の4点である⁴⁾。

聞き取った職務特性を以下に説明する(図表1参照)。事務職には、総務や経理、情報部門などの在籍者が含まれる。新たなシステム導入などは別にして、決められた事務処理を確実に遂行する

図表1：職種別の職務特性（聞き取り調査のまとめ）

	成果把握の容易さ		努力把握の容易さ	
	成果分布	測りやすさ	観察容易さ	基準の明確さ
事務職	狭い	困難	容易	不明確
工務職	広い	容易	容易	
営業職	広い	容易	困難	
生産職	狭い	困難	容易	不明確
設計職	ややひろい	やや困難	やや容易	
部門管理者	広い	容易	困難	

2) 聞き取りの実施日程については付表1に示した。

3) E社の部門管理者は、課レベルの組織単位の長を意味する。

4) 質問に当たっては、質問票を事前に準備したが、聞き取り時の話の流れにしたがって、柔軟に質問内容を取り替え、成果把握の困難さと努力観察の困難さを把握した。

ことが求められる。業務をミス無く素早く遂行することが成果である。担当業務が決まれば、ミスが出ないかぎりは、誰がやっても同じような結果が得られるので、上司が成果の個人差を測ることは容易ではない。担当者も「同じ仕事を続けており特に改善できることもないので、AやCの査定は考えにくい」という。いっぽう、同じ場所で多くの時間いっしょに仕事をするため、努力の観察は容易である。「一人ひとりが別々の仕事をしているが、電話の様子を見ていれば、順調に仕事が進んでいるのかどうかは自然に理解できる」と上司がいう。また、職場のメンバーもお互いの観察が容易であり、自分の努力を相対的に把握することも容易である。ただし、上司も部下も「A、B、Cの判断基準がわかりにくい」というので、努力過程の個人差を査定ランクに当てはめることは容易ではないと考えられる。

工務職は工事を仕切る仕事である。具体的業務には、資材と施工業者の手配や現場での工事監督などがある。工事の施工タイミングに合わせて資材や業者を手配する必要があり、過去の経験をもとにした複雑な判断力や交渉力が求められる。担当工事は、部下の能力や案件の重要さ、他案件の進捗度などを考慮して、上司が決定する。成果には、達成工事金額やその重要度、納期など、計数的に把握できるものと、工事の出来映えなどの質的なものがある。前者は成果の分布幅も広く、成果の個人差を測定することが容易であり、ともすればそれに頼って評価しがちである。しかし、上司は「売上高や利益だけでなく、工事の出来映えを加味して評価を行う」とむしろ、評価にあたって計数的な指標に頼りすぎないことを注意している。いっぽう、努力の観察は比較的容易である。なぜなら、工務職は工事監督時以外には社内に在籍していることが多いからである。案件を巡ってグループ内や営業担当者とのやりとりも多い。また上司とも頻繁に打ち合わせがある。「部下の仕事がうまく進んでいるかどうかは、職場での動きでわかるし、工程が予定通りに進んでいるかどうかでもよくわかる」と管理者がいう。また、部下も「仕事がスムーズに進んでいるかどうかをお互いに良く理解できる」という。

営業職には、継続取引の活動と新規顧客を開拓する活動の2種類がある。内容はどちらも似ている。担当クライアントのニーズを聞き取り、社内外の関係者と打ち合わせて、デザインを提案したり、見積りをしたりして、受注に結びつける活動である。売上予算は、担当者の能力と部門の営業目標を考え合わせて、管理者が一人ひとりに割り振る。聞き取った営業グループの場合、売上目標はAさんが30百万円、Bさんは250百万円、Cさんは140百万円、Dさんは220百万円、Eさんは160百万円という割り振りであった。「Aさんは、受注基盤を開拓する役割をもち、経験の少ないDさんがその受注基盤で売上をあげるという考え方で割り振っている」と上司がいう。成果として重視されるのは売上予算の達成度と次期の継続受注高である。すなわち、成果を計数的に把握できるため、その分布幅は広く、成果の個人差を測定することも容易である。いっぽう、努力の観察には困難さがある。なぜなら、「在社比率は、私は6割、部下は4割程度だと思う」と管理者が述べるように、部下が仕事を行う営業現場を上司はあまり観察できないからである。

生産職は指示された部材を工場で生産する。作業はチームで行われる。「工場はチームプレーで

運営しているので、全員に同じ目標を持たせている。また、仕事を管理者が配分するので出すべき成果は仕事配分と同時に決まる。このため、格別良い評価や悪い評価はつけない。団体生活ができない人はいない」と上司はいう。このような運営をしているため、個人成果の分布幅は小さくまた個人差の把握も難しい。聞き取り対象者となった上司は、成果の個人差を把握することに興味がない様子であった。いっぽう、努力過程は比較的容易に観察できる。なぜなら、工場では、8:30-17:00の時間、ずっと一緒に働くからである。管理者は最低でも1日に2回は現場を巡回し、「部下の顔色の悪いときなどは医者に行くことを指示したりする」と、密接な監督が行われている。また、担当者も、「一緒に作業をするので、同じ持ち場であれば、さぼっていればお互いにわかる」という。

設計職には、受注を目的とする提案設計と受注後の施工設計の2種の仕事がある。提案設計は、クライアントと打ち合わせてニーズを確認し、パース図やイラストなどさまざまな方法でイメージを具体的に表現し、提案を行うのが具体的な業務である。施工設計は工事のための図面を作成する仕事である。案件に関して調査し、作図して工務や施工業者、クライアントなどと打ち合わせるのが、その業務内容である。どちらの仕事も担当案件は、管理者が決定する。成果のうち、営業目標と関連した数値目標（販売目標やコンペ勝率など）の達成度は幅広くばらついており、成果測定は容易である。ただし、「できあがりの質については、お客さんの評判や進行写真、現地視察などによって判断する。しかし、質については部下と考えの違いが出てくるので、評価者が評価されているような気がする」と、質的な成果の測定は容易ではなく管理者が一定の負担も感じているようだ。いっぽう、努力過程については、7割程度の時間を社内で過ごすため、観察が容易である。このため、「同じ職場のメンバーで勤務態度に順位をつければ、結論は一致すると思う」と担当者が述べる。管理者も「行動予定表や案件ごとの管理表、日報、個別の報告等で状況はつかめる」という。

部門管理者は他社における課長と考えてよい。職種はさまざまであるが、共通して重要なのは、案件を担当者に配分する判断である。E社では部門単位に損益が管理されるので、予算達成度や他の重要な成果は部門業績として把握され、成果の観察は容易、かつ成果の分布幅も広い。いっぽう、努力の観察は、必ずしも容易ではない。なぜなら、仕事の割り振りのうまさや部下へのアドバイスの善し悪しを上司が直接観察することは困難であるからであり、部門の業績から間接的に推し量ることが多くなりがちである。

4-2. アンケート調査による把握

次に、聞き取り調査を補強するため、簡易なアンケート調査を実施した。調査は2005年12月～2006年2月の間に行われた。質問票の配布と回収は、社員全員に対する人事評価研修時に行われた。対象者の87%から回答を得ることができた。質問票は10項目で構成されている。そのなかに、「あなたの仕事では、上司が個人的な成果を容易に見分けることはできますか」という質問と「あなたの仕事では、あなたの努力を上司が観察することはできますか」という質問が含まれ、各々について「1：極めて困難、2：1と3の間、3：普通、4：3と5の間、5：極めて容易」の5段階の

図表 2：職種別の職務特性（アンケート調査のまとめ）

職種	成果測定の容易さ			努力観察の容易さ		
	対象数	平均値	標準偏差	対象数	平均値	標準偏差
事務	25	2.60	1.000	25	2.76	0.970
工務	46	2.83	1.122	46	2.72	1.167
営業	82	3.09	0.878	85	3.01	0.945
生産	42	2.86	1.026	41	3.32	1.171
設計	40	2.50	0.906	40	2.75	1.006
部門管理者	12	3.08	0.793	12	2.58	0.669
その他	44	2.84	1.010	44	2.80	0.978
合計	291	2.85	0.984	293	2.90	1.034

選択肢がもうけられた。得られた結果は図表 2 の通りである。

アンケート調査では、成果測定が容易とされたのは、営業職と部門管理者であり、それが困難なのは事務職と設計職である。いっぽう、努力の観察については、営業職と生産職が容易とされ、それが困難なのは部門管理者であることがわかった。

4-3. 職種別の職務特性

前 2 項で得られた職種別の職務特性には、一致する部分と一致しない部分とがある。この原因には、聞きとり対象者が、自分の仕事を中心に答えてしまいその職種の全体を捉えきれない回答をしている可能性が考えられる。そこで、前 2 項の結果が一致するものを暫定的に確かな職務特性と考え、一致しないものは職務特性が不確かであるとして考察を進める。それを整理すれば次の通りである。

- 1) 事務職は、成果測定が困難である。努力の観察については、明確な結論が得られていない。
- 2) 工務職は 聞き取り調査では成果測定も努力観察も容易とされたが、アンケートでは必ずしもそうではない。したがって、職務特性を安易に結論づけることはできない。
- 3) 営業職は、成果測定が容易であるが、努力観察については聞き取り調査とアンケート調査の結論が一致しないので、努力観察の容易さについての結論は保留する方がよい。
- 4) 生産職は、成果測定が困難であるが、努力の観察は容易である。
- 5) 設計職は、成果測定は困難である。努力観察については、容易とも困難ともいい難い。
- 6) 部門管理者は、成果測定は容易だが、努力の観察は困難である。

上記の職務特性を前提におくと、本稿の仮説は次のように具体化される。業績評価の自己評価と上司評価の相関係数が大きくなるのは営業職と部門管理者であり、それが小さくなるのは、事務職や生産職であろう。したがって、事務職と生産職については分布が歪み、営業職や部門管理者は分

布の歪みが小さいだろう。行動評価については、相関係数が大きくなるのは、生産職であり、小さくなるのは部門管理職であろう。したがって、部門管理者については分布が歪み、生産職については分布の歪みが小さくなるだろう。以上の仮説を次節で確かめる。

5. 査定データの分析

5-1. データ

本節で用いるデータは、E社から研究用に貸与された、全社員の2004年度上期の人事評価に関連するデータである。それには従業員番号、性別、学歴、部門、職種、勤続、年齢、等級と段階ごとの人事評価結果（自己評価、上司評価、合計評価、1次評価、2次評価、最終評価）が含まれる。このデータセットは、評価結果が従業員に公開された第1回目のデータであることに留意して欲しい。つまり、データセットに含まれる自己評価の結果は、前期の査定に影響を受けていないデータであり、自己評価と上司評価のギャップを純粹に示していると考えられる。

職種別の基本統計量は図表3の通りである。各段階の評価結果を見ると、事務職では業績評価、行動評価とも上司評価が自己評価を下回る。1次評価、2次評価の値は上司評価とほぼ同じであるが、最終段階で下方に修正されて、バラツキも大きく拡大した形で査定が決定している。工務職では、ふたつの評価とも上司評価が自己評価を上回る。その後2次評価、最終評価と評価が下方に修正された上で査定結果が確定している。営業職も工務職と同様の動きを示す。また、営業職では、ふたつの上司評価が1次評価に合算される段階でも何らかの操作が加えられて評価が下方に修正されている⁵⁾。生産職では、上司評価が自己評価を上回る。その後2次評価、最終評価とも大きな修正は加えられていない。設計職では自己評価と上司評価はほぼ同じである。しかしながら、2次評価、最終評価と下方に修正されつつ、バラツキは徐々に縮小している。部門管理者は上司評価が自己評価を上回る。しかし、2次評価では下方に、最終評価では上方に評価結果が修正される。

以上のように、評価の調整は職種ごとに異なった方向に行われている。しかし、最終段階では、生産職と部門管理職以外は全ての職種において下方に修正されており、人事部門の介入があったことが推測できる。人事部に確認したところ、「査定が寛大になったため、最終段階で査定を下げさせた」とのことであり、管理者が評価結果の公開を意識して寛大な評価を行った可能性が示唆されているといっていよう。

5-2. 自己評価と上司評価の相関

次に、各職種の自己評価と上司評価の相関係数を確認しておこう(図表4参照)。まず、職種によって相関係数が大きく異なっていることがわかる。業績評価の相関係数が大きい順に職種を並べる

5) 人事部が定めたルールでは、この段階は算式によって機械的に結果が導かれることになっている。しかしながら、運用のなかで合算の段階で評価結果に操作が加えられ機械的には算出されない合算結果となっている。この点は人事部に確認を行った。

図表 3：職種別の基本統計量

変数	事務			工務			営業			生産		
	対象数	平均値	標準偏差	対象数	平均値	標準偏差	対象数	平均値	標準偏差	対象数	平均値	標準偏差
性別	26	0.31	0.471	76	0.04	0.196	89	0.02	0.149	45	0.02	0.149
等級	26	5.77	1.366	76	5.43	0.957	88	6.16	1.038	45	5.47	0.944
年齢	26	45.88	8.964	76	45.28	9.619	89	44.76	8.855	45	49.82	8.100
勤続	26	23.73	9.569	76	24.79	11.114	89	24.17	10.568	45	32.24	8.550
等級滞留	26	5.85	4.567	76	6.59	5.473	89	5.46	4.843	45	7.71	5.333
学歴	26	3.04	0.871	76	2.66	0.917	88	2.76	0.935	45	2.51	0.626
業績自己評価	21	4.05	0.865	66	3.92	0.730	77	4.01	0.881	45	3.80	0.786
行動自己評価	21	4.29	0.956	66	4.09	0.717	77	4.34	0.754	45	3.80	0.726
業績上司評価	24	3.96	0.806	76	4.25	0.733	89	4.42	0.823	45	4.07	0.863
行動上司評価	24	3.88	0.900	76	4.41	0.636	89	4.63	0.774	45	4.11	0.611
一次評価	24	3.96	0.806	76	4.32	0.716	89	4.44	0.865	45	4.11	0.804
二次評価	23	3.96	0.767	74	4.27	0.782	84	4.40	0.852	45	4.09	0.821
最終評価	26	3.77	1.107	76	4.20	0.800	89	4.16	0.940	45	4.11	0.804

変数	設計			部門管理者			合計		
	対象数	平均値	標準偏差	対象数	平均値	標準偏差	対象数	平均値	標準偏差
性別	44	0.09	0.291	51	0.00	0.000	331	0.05	0.227
等級	44	6.20	1.069	51	7.22	0.415	330	6.04	1.135
年齢	44	46.11	9.366	51	49.69	4.954	331	46.60	8.729
勤続	44	25.00	10.058	51	29.55	6.807	331	26.31	10.183
等級滞留	44	6.89	5.903	51	5.20	2.850	331	6.21	4.993
学歴	44	3.07	1.021	51	2.67	0.931	330	2.75	0.912
業績自己評価	39	4.59	1.093	34	4.18	0.716	282	4.06	0.873
行動自己評価	39	4.77	0.959	34	4.47	0.706	282	4.27	0.829
業績上司評価	44	4.68	0.857	51	4.37	0.720	329	4.33	0.816
行動上司評価	44	4.68	0.639	51	4.55	0.610	329	4.45	0.727
一次評価	44	4.68	0.857	51	4.45	0.702	329	4.36	0.812
二次評価	44	4.59	0.844	48	4.38	0.866	318	4.32	0.839
最終評価	44	4.48	0.821	51	4.41	0.753	331	4.21	0.876

注1：性別は男性が0、女性が1としている。

注2：等級は1等級から10等級までである。7等級以上が管理職である。

注3：学歴は中卒を1、高卒を2、高専・短大卒を3、大卒以上を4としている。

注4：査定段階は上位から順に7,6...1という数値を割り当てた

図表 4：自己評価と上司評価の相関係数

職種	対象数	相関係数	
		業績評価	行動評価
事務	21	0.594	0.459
工務	66	0.766	0.639
営業	77	0.630	0.486
生産	45	0.522	0.666
設計	39	0.599	0.522
部門管理者	34	0.713	0.564

と、1) 工務職→2) 部門管理者→3) 営業職→4) 設計職→5) 事務職→6) 生産職、の順になる。上位の部門管理者、営業職は成果測定が容易とされた職種であり、下位の事務職、生産職は成果測定が困難とされた職種であることを思い出して欲しい。つまり、職務特性を明確にしなかった、設計職、工務職を除き、成果測定が困難な場合には自己評価と上司評価の相関係数は小さく、成果測定が容易な職種ではそれが大きいという仮説とはほぼ整合的な結果が得られた。

次に行動評価について見てみよう。まず、業績評価に比べると、生産職以外はどの職種も相関係数が小さくなっている。業績評価における計数的成果のような、上司から見ても部下から見ても一致する成果指標が存在しないことが、相関係数が小さくなる原因と考えることができるかもしれない。次に行動評価も相関係数の順に職種を並べると、1) 生産職→2) 工務職→3) 部門管理者→4) 設計職→5) 営業職→6) 事務職、の順になる。上位の生産職は努力の観察が容易な職種、下位の事務職は努力の観察がやや困難な職種である。行動評価についても、仮説と整合的な結果を得たといつてよいだろう⁶⁾。

5-3. 上司評価の特徴

次に、上司評価の分布を図表5によって検討してみよう。紙幅の都合で、4節で仮説と整合的な結果が得られた職種に限定して説明する。事務職では変動係数0.20である。また尖度は4.18と大きな値を示すが、歪度は0.07と小さい値を示す。平均値が3.96なので、評価結果が中心化する傾向が

図表5：上司評価の分布

	業績評価							
	対象	平均値	標準偏差	変動係数	最小値	最大値	歪度	尖度
事務	24	3.96	0.81	0.20	2	6	0.07	4.18
工務	76	4.25	0.73	0.17	2	6	-0.22	3.38
営業	89	4.42	0.82	0.19	1	6	-0.90	5.09
生産	45	4.07	0.86	0.21	2	5	-0.34	1.95
設計	44	4.68	0.86	0.18	3	6	-0.02	2.29
部門管理者	51	4.37	0.72	0.16	3	6	-0.68	2.22
	行動評価							
	対象	平均値	標準偏差	変動係数	最小値	最大値	歪度	尖度
事務	24	3.88	0.90	0.23	2	6	0.25	2.99
工務	76	4.41	0.64	0.14	3	6	0.04	2.76
営業	89	4.63	0.77	0.17	2	6	-0.44	3.63
生産	45	4.11	0.61	0.15	3	5	-0.05	2.67
設計	44	4.68	0.64	0.14	3	6	-0.17	2.94
部門管理者	51	4.55	0.61	0.13	3	6	-0.45	2.78

6) 職務特性の確定をアンケート調査データだけにもとづいて判断した場合には、工務職は行動観察が容易な職種、営業職は行動観察が困難な職種とされ、工務職の相関係数は大きく営業職の相関係数が小さいという仮説と整合的な結論が得られる。

見られ、相関係数が小さければ上司評価が歪んで中心化するという仮説と整合的な結果が得られたといえる。営業職は、変動係数0.19と小さな値を示す。歪度は-0.90と負方向の最大値をしめし、尖度5.09と最大値を示す。つまり多くの人は高い評価を受けながら、少数の人が低い査定を受けていることがわかる。この結果は、相関係数が大きければ分布の歪みが少ないであろうという仮説とは必ずしも整合していない。次に、生産職では平均値4.07である。変動係数は0.21と大きな値を示す。また、尖度は1.95と小さな値を示し、分布幅が広がっていることが分かる。それなりの個人差を上司がつけているようである。管理者への聞きとりでは成果の個人差を測定することに興味がないようであったが、実際には、個人差がついている。聞き取り調査の結論から見れば予想外の結果となっている。確認はできなかったが、管理者が職場リーダーを育てるための意図的な評価を行っている可能性も考えられる。また部門管理者の平均値は4.37である。歪度は-.68であり、尖度は2.22なので、寛大化の傾向をもちつつも、幅の広い分布をしている。

以上を整理すると、相関係数が大きい部門管理者については広がりのある分布が見られ、中心化する傾向が見られないので、仮説と整合的な結果を得たといえるだろう。営業職においては広がりのある分布が見られるものの、寛大化や中心化の傾向が見られ、仮説とは一部整合しない結果が得られた。また、相関係数が小さい事務職では中心化傾向が見られ、仮説と整合的であるが、生産職においてそれは見られず、仮説と異なる結果が得られた。

次に行動評価を見てみよう。まず、事務職を除けば変動係数が業績評価に比べて小さくなり、バラツキが縮小していることがわかる。また、職種間の歪度の差が業績評価に比べて小さくなるとともに、事務職・設計職以外ではその絶対値も小さくなる。尖度を見ると、3を中心に職種間の差が小さくなる。つまり、業績評価に比較すれば、分布はより正規分布に近づき、職種ごとの特徴が不鮮明になる。

職種別にみると、生産職では、平均値が4.11でかつ歪度、尖度、変動係数も小さいので分布は正規分布にやや近い。ただし、レンジが小さいためにやや中心に分布が集まる。分布の形からすれば比較的歪みが少なく、仮説と整合的な結果が得られたといってもよいだろう。事務職は平均値が3.88となり、歪度が0.25となっているので分布の重心はやや低い目である。変動係数を見ると、比較的大きな値を示すが、尖度を見れば比較的正規分布に近い分布となる。この結果は、必ずしも仮説と整合的とはいえない結果であった。

以上をまとめると、仮説と整合的な結果が得られたのは、業績評価においては、部門管理者、事務職であった。営業職では仮説と一部が整合的、生産職では非整合的な結果が得られた。行動評価においては、生産職では仮説と整合的な結果が得られ、事務職においては仮説と整合しない結果が得られた⁷⁾。

7) 工務職について自己評価と上司評価の相関係数だけに着目すると、業績評価、行動評価の両方とも大きな相関係数を示す。上司評価の分布は、業績評価は上方に重みを持った分布となっているが、極端な中心化は見られないがゆえに、仮説と部分的に整合していると考えてよいだろう。行動評価については上方に重みがあるものの左右対称の分布となっており、仮説と整合的な結果が得られたと考えられる。つまり、工務職については聞き取り調査結果のみが仮説と異なることになった。

まとめ

本稿では、モニタリングの負担が自己評価と上司評価の相関に与える影響と、結果としての評価の歪みを企業内の人事マイクロ・データとアンケート調査、聞き取り調査を併用することによって明らかにした。得られた結論は、業績評価については事務職、部門管理者について仮説と整合的、営業職については仮説と部分的に整合する結果が得られた。行動評価については、全体的に業績評価に比べて査定の幅が縮まり、明確な結論は得られにくかったが、生産職だけは、仮説と整合的な結果を得た。総じて言えば業績評価では6つの職種のうち3つの職種が仮説と整合的であり、行動評価においては1つの職種が仮説と整合的な結果となった。全体的にはやや不明確な点もあるものの仮説とほぼ整合する結果が得られたと言ってもよいだろう。

業績評価について仮説と比較的整合し、行動評価については仮説と整合する結果が得られにくかった原因は、次のように考えることができよう。まず、昇格や昇進を意識する管理者が昇格候補者に対して有利な査定になるよう、行動評価で評価結果を操作している可能性がある。生産職の行動評価や業績評価の分布が広がるのは、管理者が意図的に格差をつけ昇格候補者を育成していると考えることによって説明できる。また、成果測定が明確な基準で行われたときに、評価全体を調整するために評価者が行動評価を操作することも考え得る。第2には、努力の観察が可能でもそれを評価基準に当てはめることが困難であることが影響している可能性がある。行動評価の標準偏差が業績評価に比べて小さくなる原因は、評価基準の曖昧さによって明確な差をつけた評価が行いにくいことにあるだろう。

聞き取り調査とアンケート調査の結果が整合的でない原因は、聞きとり対象者が、当該職種に関して偏った見方をしていることが考えられる。企業内で同じ職種と分類されていても、細かく観察すればその職務特性が異なっている可能性もある。

本稿は、部分的に仮説と整合しないという弱点があるが、それを承知したうえで、本研究が示唆することを述べたい。第1は、職務特性によって自己評価と上司評価の相関が異なることの実践的意味である。これは、面接制度によって評価に対する部下の納得性を高めるためには、努力の観察や成果測定の補助手段を必要とすることを示唆する。象徴的行動を示すとか、目標の困難度をそろえるといった評価基準整備の作業を、職種ごとに丁寧に積み重ねることが必要であろう。第2に、評価者負担の存在によって、査定に歪みが生じるという点である。梅崎・中嶋・松繁（2004）が指摘した評価者負担による査定幅の縮小という現象とも整合的であり、中嶋（2004）が指摘した評価者負担の大きさによって査定根拠が異なるという点とも整合する。従って、研究上、査定結果をパフォーマンス指標として用いることには、慎重な態度を要するであろうし、実務的には職務特性に合致した人事評価制度を構築する努力が必要であることを示しているだろう。人事評価結果が評価者を取り巻く環境によって変わりうるものであることを強く意識しておくことが必要であろう。

本研究は自己評価と上司評価の相関や職務特性に焦点をおいた端緒的研究である。それ故に残された課題は多い。第1に職務特性を把握する指標とその把握方法をより正確に探求する必要がある。聞き取り調査を通じて職務特性をとらえるための指標を開発するとともに、アンケート調査でそれを確認することを繰り返す地道な研究を蓄積する必要があるだろう。第2には、本研究は人事評価が従業員に公開された初回の査定結果を用いている。面接が積み重ねられることによって、上司評価と部下評価の一致度が高まる可能性も考えることができる。従って、面接による学習効果を把握したうえで本研究と同様の研究を行い、職務特性との関連を再度考察する必要があるだろう。最後に本研究は1社についての研究であり、対象企業の制度や人事部門の行動が上司評価に与える影響はE社独自のものである可能性もある。異なる企業においても同様の傾向が見られるかどうかはまだまだ未解明であるとすべきであろう。今後も研究の蓄積を行いたい。

謝辞：本研究はE社の多大なご協力によって可能となった。人事部担当者、聞き取り調査に対応してくださった従業員の方々に深く感謝したい。また、ともに聞き取り調査を実施した井川静恵氏（星陵短期大学講師）に深く感謝したい。同時に、本研究に的確なアドバイスをくださった松繁寿和氏（大阪大学大学院国際公共政策研究科教授）に感謝の意をささげたい。

参考文献

- 梅崎修 中嶋哲夫 (2005)、「評価者負担が評価行動に与える影響－「人事マイクロ・データ」と「アンケート調査」の統計分析」、『日本労働研究雑誌』、No.545、p.40-50
- 奥野明子 (2004)、「目標管理のコンティンジェンシー・アプローチ」、白桃書房、第6章、p.169-202
- 厚生労働省 (2005)、「平成16年度労働経済の分析」、p.228
- 厚生労働省 (2002)、「平成14年度雇用管理調査」、
<http://www.dltk.mhlw.go.jp/toukei/kouhyo/data-rou2/data14/26.xls>
- 城繁幸 (2004)、「内側から見た富士通－「成果主義」の崩壊－」、光文社
- 中嶋哲夫 松繁寿和 梅崎修 (2004)、「賃金と査定に見られる成果主義導入の効果」、『日本経済研究』、No.48、p.18-33
- 中嶋哲夫 (2004)、「評価者負担が査定におよぼす影響－営業所長の査定から」、『日本労務学会誌』、第6巻第1号、p.36-43
- 中嶋哲夫 (2007)、「人事評価制度の納得度を決定する要因－職場目標への納得度と職務配分の公正さ－」、日本労務学会誌、第9巻1号、p.68-80
- 柳下公一 (2003)「ここが違う勝組企業の成果主義：対話と個の確立を目指して」、日本経済新聞社、Milkovich, George T. and Boudreau, John W. (1997), "Human Resource Management 8th ed.", IRWIN, chapter4, p.112
- Milgrom, P. and Roberts, J (1992), "Economics, Organizations & Management", Prentice Halling., chapter12, p.388-422 (邦訳、奥野正寛・伊藤秀史・今井晴雄・西村理・八木甫、「組織の経済学」NTT出版、1997、p.433-472)

付表1：インタビュー実施日程と対象者

面接日	会場	面接時間	対象者	面接者
8月19日	大阪	9:30-10:30	事務担当者 (被評価者)	中嶋、井川
		10:40-11:40	設計担当者 (被評価者)	
8月22日	東京	9:30-10:10	営業管理者 (評価者)	中嶋、井川
		10:30-11:30	営業担当者 (被評価者)	
8月23日	東京	9:30-10:15	工務管理者 (評価者)	中嶋、井川
		10:20-11:00	工務担当者 (被評価者)	
		11:20-12:10	事務管理者 (評価者)	
8月30日	大阪	9:30-10:15	生産担当者 (被評価者)	中嶋、井川
		10:20-11:30	設計管理者 (評価者)	
		17:00-18:00	生産管理者 (評価者)	