

## 研究ノート

# 業績評価システムにおける因果連鎖情報の伝達： 文献サーベイ

西 居 豪

### 1. はじめに

BSC (balanced scorecard) の発展 (Kaplan and Norton, 1992, 1993, 1996, 2001, 2004, 2006, 2008) と普及に伴い、業績評価研究は活性化し、非常に多様なトピックについて研究が蓄積されてきている (西居, 2011)。

複数の視点とそこに配置される複数の業績評価指標、さらには成果指標とパフォーマンス・ドライバーの区分など、BSC はかなり多様な情報を包含したシステムとなっている。しかしながら、その複雑さゆえに、状況の正確な認識やそれを踏まえた業績評価や意思決定に難しさを抱えており、BSC 導入やその活用の妨げとなっているという発見がなされている (Lipe and Salterio, 2000; Malina and Selto, 2001; Ittner et al., 2003ab; Ittner and Larcker, 2003)。こうした課題への対応方法として注目されているのが、BSC の鍵概念とも言われている、戦略目標あるいは業績評価指標間の因果連鎖 (causal linkage) である。Kaplan and Norton (1996) にて、4つの視点を因果的に結びつけて理解することが強調され、Kaplan and Norton (2001, 2004) では、それが戦略マップというグラフィカルなツールに埋め込まれる形で実用化された。その結果、因果連鎖情報が提供されることで生じる影響に多くの研究関心が向けられることになった。

本研究ノートでは、業績評価システムにおける因果連鎖と評価バイアスとの

関係性について検討した先行研究<sup>1</sup>の文献サーベイを通じて、どういった問題に注目が集まり、何が明らかにされているのか整理し、今後の研究の方向性を提示することにした。

次節以降の構成は以下の通りである。2節では、BSCの提唱者であるKaplanらによる因果連鎖概念の説明を概観した上で、彼らの論考に対する批判についても検討する。3節では、既存研究のサーベイを通じ、問題の焦点を明らかにするとともに、何が解明されてきているのか整理する。3節の整理を受けて、4節では既存研究では十分な関心が向けられていないけれども、検討する価値が高いと思われる研究領域を提示する。最後の5節はまとめである。

## 2. 因果連鎖

本節では、因果連鎖の定義に関するKaplanらの説明とそれに対する批判的考察について簡潔にまとめておくことにしたい。

Kaplan and Norton (1996, p.149) は、戦略を原因と結果に関する一連の仮説であると定義し、この仮説が4つの視点の目標間あるいは指標間の関係に反映されるBSCを構築する必要があると主張している。こうした関係が因果関係(cause-and-effect relationship)<sup>2</sup>と呼ばれている。また、彼らの説明では、視点を横断した関係性のみならず、成果の指標、戦略的イニシアティブ、パフォーマンス・ドライバーという各視点内での因果的な繋がりについても検証することができる」と説明している。ただし、統計的な検証というのは必ずしも前提とされており、管理者にとって、その繋がりや程度あるいは有無を検討するこ

---

1 因果連鎖情報を扱った既存研究にて検討が試みられているトピックは多様ではあるけれども、紙幅の関係上、本研究ノートでは評価バイアスの議論に限定した文献サーベイを行う。

2 Kaplan and Norton (1996) の中では、causal relationship や causal chain of relationships あるいは linkage of measures across the four perspectives といった表現も同義的に用いられている。既存研究でも名称については、バラツキが見受けられるが、後でサーベイする研究では causal linkage という表現がしばしば用いられている。本研究ノートでは、因果関係という用語は後述する厳密な意味での因果性を踏まえた関係を一般的に連想させるので、次節以降では、因果連鎖という用語を用いることにする。

とは、たとえそれが非公式のものであったとしても有用であるというスタンスに立っている (Ibid., p.261)。

このように、目標あるいは指標間を結びつけるというアイデアはかなり強く主張されているけれども (Kaplan and Norton, 1996, 2001, 2004), その解釈や実装に関しては学術的にも実務的にもかなり混乱がある (Bukh and Malmi, 2005)。

学術的観点から BSC における因果関係の理論的基盤に疑問を呈したのは Nørreklit であった。彼女は、Kaplan らがいう因果関係が、独立性 (原因 X と結果 Y が独立している)、時間的順序 (原因 X が結果 Y より時間的に先行している)、予測能力 (原因 X が観察されたら必ずその後 Y が観察される) という因果性の条件を満たしておらず、論理的関係や目的論関係が混在した説明を展開していると批判している (Nørreklit, 2000)。たとえば、ROI 公式のような会計モデルの論理から導かれている要素間の結びつき (たとえば、売上高利益率と ROI の関係) は、経験的に検証されるものではなく、因果関係として位置づけることはできないという。そして、彼女は、BSC への関心の高さが、革新的で実践的な理論としての BSC の実態というよりもむしろ、宣伝用のレトリックによるところが大きいと結論づけている (Nørreklit, 2003)。

因果関係の判定条件については、議論のあるところではあるが、Nørreklit (2000) が取り上げたものは、適用されることの多い項目である (Malina et al., 2007)<sup>3</sup>。これらの条件に従えば、確かに Kaplan らの説明には、Nørreklit (2000) が指摘したような、概念定義としての曖昧さや不正確さがあると思われる。ただし、実践的な観点からは、Nørreklit の批判は BSC そのものの有効性を否定するほど妥当ではないかもしれない (Bukh and Malmi, 2005)。いくつかのケース・スタディは、厳密な意味での因果関係が組み込まれていなくても、BSC が有

3 米国公衆衛生局長諮問委員会は、上述した3条件を含む、異なる地域・時代・状況でも同一のことが生じる一貫性 (consistency)、原因 X と結果 Y の関連が強い強固性 (strength)、原因 X が変われば結果 Y も変わる特異性 (specificity)、原因 X が結果 Y に先行する時間性 (temporality)、既存の知識体系と矛盾しない整合性 (coherence) という5つの基準を提示している。

効な業績評価システムとして機能する可能性を示唆する発見を報告している。

BSCの実施に関連した実態調査の結果は、因果的な関係をBSCに反映させていなかったり、指標間関係の統計分析も実施していない企業が少なくないことを示している (Ittner et al., 2003b; Speckbacher et al., 2003)。分析をしていないわけだから、そうした企業では、BSCの指標間関係に妥当な原因と結果の関係があるのか定かではない。この点について調査を行ったのが、Malina et al. (2007)である。彼女らは、とある企業の北米販売店チャンネル向けに開発されたBSCにおける31四半期分の業績データを用いて、先行指標によって遅行指標の動きを説明する予測モデルの妥当性について実証分析を行った。その結果は、予測モデルを支持する十分な証拠は得られなかったというものである。にもかかわらず、インタビュー調査で得られた関係者らの声は、BSCの利用に満足し、さらなる拡張を計画しているというものであった。Malinaらは、これらの分析結果を踏まえて、BSCの指標間関係の因果性について統計的な証拠が欠如しているからといって、BSCの妥当性を安直に否定すべきではないと主張している。同様にWebb (2004)の実験結果やSundin et al. (2010)のフィールド・スタディは、モチベーションや目標整合性などの行動的な機能から見た場合、因果関係が実在することよりもむしろ、そうした関係が認められると組織成員が知覚していることがより重要になることを示唆している。

一方、Kober and Northcott (2021)は、公的部門のBSCのパネルデータにグレンジャーの因果性検定を適用し、当該BSCに因果関係が認められることを報告している。また、指標間関係の因果的な妥当性を統計分析によって検証したり、システム・ダイナミクスに基づいた定量的なシミュレーション・モデルを組み込んだりしている事例も観察されている (Barnabè, 2011; Epstein and Westbrook, 2001; Kaplan and Norton, 2001)。ただし、上記の因果関係の定義とも整合するランダム化比較試験を適用した因果効果の発見・推定 (伊藤, 2017; 安井, 2019; Leigh, 2018; Kohavi et al., 2020)に基づいたBSCの構築事例は報告されておらず、業績データの統計分析を通じた関係性の検証が議論の中心となっている。

個別の企業単位で見れば、構築されたBSCにおける戦略目標・指標間に因果性が認められるかどうかは、策定された戦略そのものの妥当性やその確実な実行力に大きく関係する。その意味では、BSCにおいて因果関係とされるものは、一般的なものとして認識されるべきではなく、組織、実際の状況、関連する時間軸に固有のものである（Bukh and Malmi, 2005）。さらに、正確性を最優先した情報システムではなく、組織成員の戦略に整合した行動を誘発する影響システムとしての側面が重要視されているので（Kaplan and Norton, 1996, 2001, 2004）、因果関係があるか否かというよりもむしろ、因果的な情報提示が人々の業績評価や意思決定にどのような影響を及ぼすのかという点に既存研究の関心は向けられている。次節では、こうした研究関心を有し、特に評価バイアスについて検討を行っている先行研究をサーベイすることにしたい。

### 3. 文献サーベイ

本研究ノートでレビュー対象とした文献は、下記の手続きを経てリストアップされたものである。まず、アメリカ会計学会のデジタル・ライブラリーとScience Directにて、「strategy map」, 「causal linkage」, 「causal chain」, 「causal model」がAbstractもしくはTitleに含まれている文献（1996年以降）を検索した<sup>4</sup>。なお、キーワードの特質上、自然科学分野の雑誌も含むScience directでは、ヒット件数があまりにも大きくなるので、雑誌名にAccountingが含まれていることを条件として追加している。これらの検索条件でヒットした文献は45本であった。そこから、業績評価システムにおける因果連鎖と評価バイアスを対象とした議論ではない文献を除外した。さらに、これらの文献における、上記のトピックを扱った引用文献もレビュー対象とした。以下では、こうして抽出された7つの文献の発見事項を詳細に示すことにしよう。

---

4 「cause-and-effect relationship」は一般的な因果関係としての意味合いで使われる用語なので、今回の検索キーワードからは除外した。ただし、BSCやperformance measurement systemなどの関連ワードとAND検索することで、重大な文献漏れがないか確認した。

### 3.1. 共通指標バイアス

因果連鎖に関する情報提供と評価バイアスとの関係について検証した先駆的研究は Banker et al. (2004) である。彼らは, Lipe and Salterio (2000) で発見された共通指標バイアスを緩和する方法として, 戦略に関する情報提供に注目した<sup>5</sup>。共通指標バイアスとは, BSCのフォーマットを用いて, 複数のビジネス・ユニットの業績を評価する際に, 評価者が認知負荷の観点で単純化戦略に従ってしまうことで, 異なるビジネス・ユニットに共通した指標を重視し, 戦略的に重要である可能性の高い個々の事業に特有な独自指標を軽視・無視してしまうバイアスである。

Banker らの実験では, Lipe and Salterio (2000) で設定された共通指標と独自指標という分類に, 指標の戦略とのリンクの有無という条件を付け加えた。すなわち, 戦略にリンクした共通指標, 戦略にリンクしない共通指標, 戦略にリンクした独自指標, 戦略にリンクしない独自指標という4種類の指標が, 4つの視点にそれぞれ配置されるBSC(指標の数は全部で16)が設定された。2つのビジネス・ユニットの業績を評価するタスクであるので, 4種類の指標の測定実績にて, いずれのユニットが優れているのかどうかによって, 計16のBSC(実績値込み)が作成された。この16パターンのBSCに対して, 480名のMBAの学生が30名ずつ割り当てられた。そして, 30名の半分の15名には, 企業とユニットの簡潔な情報のみが提供されたのに対して, 残りの半分には, それらの簡潔な情報に加えて戦略の記述情報と戦略マップも提供された。被験者は, 上記のBSCの実績データに基づき, 2つのユニット長の業績を評価するよう求められた。

ユニット長の業績評価スコアの差に注目した分析から, 戦略に関する詳細な情報が提供される場合には, 戦略にリンクした指標をより重視した評価が行わ

---

5 共通指標バイアスを緩和させる方法は他にも発見されており, 評価プロセスの説明責任の付与あるいは第三者による保証報告の提供 (Libby et al., 2004), 評価プロセスの分解 (Roberts et al., 2004), BSCに関する教育訓練 (Dilla and Steinbart, 2005), 複数回の成果のフィードバック (Krumwiede et al., 2013) などがある。

れること、戦略にリンクしない共通指標よりも戦略にリンクした独自指標をより重視した評価が行われることを明らかにした。つまり、この結果は、戦略に関する詳細な情報提供によって、共通指標バイアスの問題が解決されることを示唆している。

Banker はさらに、Banker et al. (2011) にて、Banker et al. (2004) で得られた知見を拡張した。注目されたのは、戦略に関する物語情報を戦略マップで補うことで、戦略目標により合致した業績評価に繋がるのかという疑問である。戦略にリンクした指標の2つのユニットの相対的業績、戦略にリンクしていない指標の2つのユニットの相対的業績、戦略情報の提供（①会社の概要を示す文章のみ、②各ユニットの戦略を説明する簡単な物語情報（①に加えて）、③4つの視点に配置された戦略目標を矢印で結んだ各ユニットの戦略マップ（②に加えて））の3要因が被験者間要因とされた。こうしたトリートメント設定であるので、マップの提示がナラティブな情報提示よりも優れているかどうかではなく、むしろ戦略マップという補助的なグラフィカル表示によって、ナラティブな情報提示のみの場合よりも、BSCの効果的な利用が促進されるのかどうか、という点に研究関心が向けられている。なお、各ユニットに固有の戦略マップ情報はすべて物語情報から推測できるという点で、③で追加提示される戦略マップは②で提示される物語と情報的に等価になるように設計されている。

2つのユニット長の業績を評価した際の被験者ごとのスコア差を従属変数とした回帰分析の結果に基づいて、各トリートメント群における戦略にリンクした指標とリンクしていない指標の平均的な影響が計算された。その結果、①や②と比べて、③の条件である戦略マップも提供された場合には、戦略にリンクした指標がより重要視されることが明らかとなった。Banker et al. (2011) は、Banker et al. (2004) にて記述情報と戦略マップがセットでトリートメントされていたものを分解しているため、得られた分析結果は、戦略マップというグラフィカルな情報の追加提示によって、BSCを用いて業績評価する際の認知的な困難さが低減されることを示唆している。



Humphreys and Trotman (2011) は、Banker et al. (2004) で検討された条件をより細かく捉えた。Humphreys らが特に注目したのが、本来の BSC であれば、利用される指標の全てが戦略とリンクしていること (Kaplan and Norton, 1996) が求められるにもかかわらず、Banker et al. (2004) では、半分の指標しか戦略とリンクしていなかったということ<sup>6</sup>である。つまり、Banker らの実験設定は、教科書的には決して望ましいとは言えない、戦略とリンクした指標とリンクしていない指標とが BSC の中に混在している状態であった。こうした指標の組み合わせは、業績評価のタスクを完遂するのに必要な認知努力水準を高めてしまうのに対して、逆にすべての指標が戦略とリンクしているなら、こうした評価タスクの認知上の複雑性を減らすことが期待できる。こうした予測のもと、Humphreys らは、戦略情報の提供と戦略にリンクした指標によって、共通指標バイアスを取り除くことができるのがどのような場合なのか実験した。

戦略と指標のリンク (BSC の半分・すべて)、戦略情報の提供 (戦略の記述情報と戦略目標間の因果連鎖を描写した戦略マップの提供の有無)、評価対象となる2つの事業部の業績優劣パターン (共通指標・独自指標のいずれかで優勢) の3要因を被験者間要因とし、これら要因による2人の事業部長の業績を評価したスコア差への影響を分析した。得られた結果は、指標の半分しか戦略とリンクしていないときは、戦略情報提供の有無にかかわらず、独自指標に比べて共通指標により大きなウェイトをおいた評価をする共通指標バイアスが生じてしまうことが明らかとなった。さらに、被験者間要因の1つである戦略と指標のリンクを、戦略マップにおけるテーマの違いに入れ替えた実験により、すべての指標が戦略とリンクしている状況では、戦略情報の提供によって、共通指標バイアスを除去できることが分かった。逆に、戦略テーマが違ってても、戦略情報が提供されない場合には、共通指標バイアスの発生が確認された。つまり、戦略情報の提供のみならず、すべての指標が戦略とリンクしていなけれ

---

6 戦略にリンクしていない共通指標・独自指標が各視点に1つずつ配置されているので、16ある指標のちょうど半分が戦略とリンクしていない指標となる。



ば、共通指標バイアスが緩和されないことが明らかにされた。この結果は、オペレーショナルな非戦略的な指標をBSCに組み込むことが望ましくないという、BSCの基本的な特徴の重要性を示唆している。

### 3.2. 原因の帰属に関連したバイアス

共通指標バイアスは戦略マップ以外の緩和手段も含めれば、かなり研究蓄積が進んだトピックであるが、他のバイアスも検討対象になっている。たとえば、複数の評価対象ではなく、単一の管理者の業績や戦略の評価という状況下でも、偏った評価行動は観察されている。

Rotaru et al. (2020) は、因果的マルコフ条件に違反する形で結論を出すという因果推論におけるバイアス (Rehder, 2014) が、BSCを用いた業績評価でも発生するのか、実験を行った。このバイアスは、BSCの文脈では、因果連鎖の仲介変数に該当する2つの視点（顧客の視点と内部ビジネスプロセスの視点のことを指し、インナー視点と呼ばれる）が軽視され、根本原因（学習と成長の視点）と最終結果（財務の視点）（これら2つの視点はアウター視点と呼ばれる）を直接結びつけてしまうバイアスとして理解される。さらに、戦略マップが必ずしも組織の価値発生プロセスの複雑性を忠実に反映したものではなく、むしろ単純化して表示していることから、マップでは示されていない他の潜在的な発生原因を推測させてしまうことで、上記の因果連鎖の最初と最後を過度に強調するのではないかという疑問も検証された。

業績結果のパターン（アウター・インナーのいずれの業績結果がポジティブあるいはネガティブか）と戦略マップの提示（マップあり・マップなしで戦略目標のみ）の2要因を被験者間要因として、BSCの実績値（目標値も併記）に基づいて、管理者業績を評価する実験を行った。その結果、因果的マルコフ条件を違反し、インナー視点の業績よりもアウター視点の業績によりウェイトのおいた評価がなされる傾向があることが明らかとなった。また、戦略マップの提供によって、アウター視点の業績を過度にウェイトづけて評価するバイアスがさらに深刻化することも明らかにされた。なお、追加の分析と実験を行う

ことで、アウトナー視点の1つである財務の視点だけにウェイトがおかれているのではなく、学習と成長の視点も含めたアウトナーの2つの視点両方をより強調した評価を行っていることが確認されている。Rotaru らはこの結果を受けて、戦略マップの提供は共通指標バイアスへの対処法として有用なツールである一方で、BSC にて因果的に構成される要素の歪んだ解釈を誘発してしまう危険があることを指摘している。

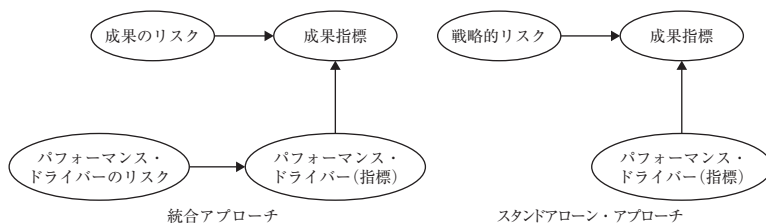
Cheng et al. (2018) は、根本原因をより重要視するという人の傾向 (Yopchik and Kim, 2009) に注目し、BSC による戦略評価にリスク情報が及ぼす影響について実験を行った。競合企業の行動の変化、内部プロセスの重要なエラー、技術的ケイパビリティの損失といった、意図した戦略の遂行を阻害する意図せざる出来事である戦略的なリスクに関する情報は、戦略を評価対象とするBSCの運用を考える上で重要である。しかしながら、そうしたリスク情報をBSCに取り込むことで戦略の評価にどのような影響をもたらすのかは未解明なままとなっているというのが、彼らの問題意識である。

具体的に注目されたのはリスク情報のプロフィールと表示方法の2つである。まず、プロフィールとは、以下のように操作化された。リスク要因が運行的な財務と顧客の2つの視点（成果と呼ばれる）あるいは先行的な内部ビジネス・プロセスと学習と成長の2つの視点（パフォーマンス・ドライバーと呼ばれる<sup>7</sup>）のいずれに関連するものか分類した上で、リスクの高さ（生起確率が1～10の得点で示されている）が相対的に成果とパフォーマンス・ドライバーのいずれで高いかによって2つのプロフィールに区分される。次に、表示方法は、これらリスク情報がBSCに追加の欄を設けて示す統合アプローチと別々の表の形で示すスタンドアローン・アプローチに区分される。統合アプローチでは各視点に関連するリスク情報が割り当てられるのに対して、スタンドア

---

7 成果指標とパフォーマンス・ドライバーという名称は、各視点内の戦略目標の進捗度を測定する指標と目標達成のためのイニシアティブの実行度を測定する指標を示す、いわゆる横の因果連鎖について言及するときに使われるものであるが (Kaplan and Norton, 1996, 2001), Cheng et al. (2018) では、こうした意味で用いられていないので注意が必要である。

図1. 統合アプローチとスタンドアロン・アプローチ



(出所：Cheng et al., 2018, p.97)

ローン・アプローチでは、戦略への脅威として一括りに示されており、特定の視点との関連づけは行われていない（図1参照）。人に根本的な原因をより重要視する傾向があるのであれば、根本的な原因をより強調して示す統合アプローチのもとでは、成果のリスクがより高い戦略よりも、パフォーマンス・ドライバーのリスクがより高い戦略を否定的に捉えるのかどうか注目された。

上記のリスクに関する2要因が被験者間要因とされたが、統制群として何もリスク情報を受け取らない群が設定された。そして、12か月前に遂行された戦略の業績を調査する新しく任命された管理者の役割を果たすよう求められた被験者の評価スコア<sup>8</sup>を従属変数とした分析が行われた。結果は次の通りである。

統合アプローチのもとでは、戦略が成果よりもパフォーマンス・ドライバーに関して相対的に高いリスクを抱えているときに、管理者は他のビジネス・ユニットに対して戦略拡張の推薦をしない傾向がある。反対に、戦略的リスクがスタンドアロンなリスト形式で示されているときには、管理者は異なる戦略的リスク・プロファイルの区別をしていない。

戦略の評価についても、拡張推薦の評価と同様の傾向が見られた（統合アプローチのもとでは、高いリスクが成果ではなくパフォーマンス・ドライバーに関連している場合に、好ましくない評価を下す）けれども、リスク情報提示アプローチと戦略的リスク・プロファイルとの有意な交互作用効果は観察されな

8 戦略の業績を評価したスコアと他地域で当該戦略が遂行されるべきかどうか推薦する可能性を評価したスコアの2変数である。

かった。

仮説としての戦略が組み込まれたBSCによる業績評価では、戦略の妥当性を毀損する可能性のある情報の検討が欠かせない。戦略的な適応に関して、サーベイ研究では、BSCのインタラクティブな利用が頻繁に検討されているが（たとえば、Bisbe and Otley, 2004; Speklé et al., 2017）、コミュニケーションによる情報共有に注目が集まっており、リスク情報を個人がどのように認知するのかという点は検討されていない。個人のリスク情報認知は戦略的適応行動の起点になり得るので、Cheng et al. (2018)の知見はサーベイ研究の発見を補完するものとして高く評価できるだろう<sup>9</sup>。

Taylor (2010) は、動機づけられた推論による戦略（イニシアティブの効果）の偏った評価に注目した。動機づけられた推論とは、自分が信じたい結論を支持する論拠は、信じたくない結論を支持する論拠よりも強いと感じる傾向（Kunda, 1990）である。特に、Taylorが注目したのは、イニシアティブを自ら選択した場合に、動機づけられた推論によって、楽観的な戦略評価がなされてしまうことであった。こうした歪んだ評価が下されるのかどうか、そしてこれを緩和するのに、因果連鎖に関する情報提供とBSCで用いられる指標の選択への関与が効果的なのか検証した。

因果連鎖に関する情報提供（BSCの4視点による分類、視点間に矢印を引いたマップ）とスコアカード実行への関与度（非関与、イニシアティブ選択への関与、イニシアティブとその成功度を測定する指標の選択への関与）の2要因を被験者間要因として、前年度データからの変化率と実績値が掲載されたBSCをレビューし、イニシアティブの成功度を評価する実験を行った。その結果、イニシアティブの選択に関わった被験者は、自ら選択することなくトッ

9 本研究ノートでは、紙幅の関係上、詳しく取り上げることはできなかったけれども、同様の観点からの実験研究としてCheng and Humphreys (2012)がある。彼らは、外部情報の戦略との関連性の判断、それら外部情報に基づいた戦略の適切性の評価に対する、戦略マップによる戦略目標間の因果連鎖情報の提供とBSCの4つの視点への指標の分類の影響について検証した。その結果、戦略マップによる因果連鎖情報の提示によって、外部情報と戦略との関連性（無関連性）とそれら情報に基づいた戦略の適切性をより正しく評価できるようになることが明らかにされている。

ブが選択したと説明された被験者よりも、そのイニシアティブをより成功したものと認知する傾向があることが確認された。そして、この戦略評価タスクにおける動機づけられた推論は、因果連鎖に関する情報が提供されるだけでは克服されなかった。楽観的な戦略評価という偏りを解消するには、因果連鎖に関する情報提供に加えて、イニシアティブの評価に用いる指標の選択に自ら関わることも必要であることが明らかにされた。

Taylor (2010) の発見は、戦略実行への動機づけの観点からイニシアティブの考案・策定を下位の管理者や従業員に委ねた場合、実行者自らの戦略評価と権限を移譲した上司の戦略評価との間に対立を生む可能性を示唆している。さらに、評価指標の選択への関与によって、この歪みを解消する方法も発見できる可能性を示した。それゆえ、Taylor の発見は実務への警告と実践的処方箋の提唱という両面で興味深いと言えるだろう。

Taylor よりも明示的に、評価者と被評価者との対立構図に注目した研究が Wong-On-Wing et al. (2007) である。彼らは、評価者によるバイアスが、業績の潜在的な決定要因として、評価される管理者の努力と能力に比べて、戦略の質に十分な注意を払っていないことに起因しているのではないかという疑問の解明に取り組んだ。こうした疑問への注目を後押ししたのは、BSC 導入企業にて、評価者と被評価者との間で対立や緊張が生じたことを観察したフィールド・スタディ (Ittner and Larcker, 2003; Ittner et al., 2003b; Malina and Selto, 2001) の存在である。具体的に注目されたのは、評価者の選択的注意に関連した、行為者観察者バイアス (Jones and Nisbett, 1971) と対応バイアス (Gilbert and Malone, 1995) の2つのバイアスである。前者は、行為者として自分の行動の原因を検討する際には、その状況の影響力も考慮するにもかかわらず、観察者として他者の行動の原因を検討する際には、行為者である他者の性格や能力といった内的な特性を重視する傾向を意味する。後者は基本的帰属錯誤とも呼ばれるように、外的な状況を無視して、性格や能力といった内的な側面に原因を求める傾向を意味する。これらのバイアスが生じる傾向を踏まえると、評価者 (観察者) は、被評価者 (行為者) よりも、行為者の業績が状況的要因 (たと

えば、戦略そのものの質)によるものであると認識する可能性は低く、たとえ戦略の質が低く、業績が低迷している部分があったとしても、それをすべて行為者に原因があると捉えてしまうリスクがある。Wong-On-Wingらは、こうしたバイアスの出現によって、評価者と被評価者との間で評価をめぐる対立が生じるのか、そして、業績を決定する戦略の重要性を評価する機会を設けることで、このバイアスを緩和できるのか、実験を行った<sup>10</sup>。

被験者の果たす役割(トップマネジメントと店長)と評価機会(店長の業績の評価前に当該業績に対する戦略の重要性を評価する機会の有無)の2要因を被験者間要因とした、買取された2つの店舗の店長の業績(BSCで測定)を評価することを被験者に求めた実験を行った。分析の結果、トップマネジメントの役割を担う被験者は店長の役割を担う被験者よりも店長の業績をより低く評価する傾向があること、このバイアスは評価者であるトップマネジメントに業績に対する戦略の影響を考慮させる仕組みを設けることで解消できることが明らかにされた。

戦略にリンクした指標を重要視した評価を行えるかどうかは重要であるけれども(Banker et al., 2004)、Wong-On-Wingらの実験結果は、たとえ戦略にリンクした指標を重要視していたとしても、管理者の業績を評価する際に、必ずしも戦略そのものの有効性による影響を加味するとは限らないことを示唆している点で、BSCの設計・運用に対して、重要な含意を含んだ発見と言えるだろう。

#### 4. 今後の研究課題の方向性

本節では、今後の研究にて検討の価値が高いと思われる領域を提示したい。

---

10 評価者と被評価者との間の対立に関しては、明示的にバイアスについて言及しているわけではないが、Johnson et al. (2014)の研究もまた興味深い。彼らは、BSCのようなタイムラグを伴う視点間の因果連鎖を想定した業績評価システムにおいては、被評価者がまだ責任を負えない将来の財務業績を評価者が過度に重視してしまう可能性があることを実証している。さらに、管理不能で責任を負うのが適切とは言えない財務指標に対する評価者の固執を減少させるのに、戦略マップの各視点において、明確な戦略実行のタイムライン情報を提供することが効果的であることも明らかにしている。

#### 4.1. 主観的評価

評価に関するバイアスの発見とそれへの対処方法の提案は理論的貢献のみならず実践的含意も有するものとして高く評価できるだろう。しかしながら、BSCの4つの視点に配置された複数の業績評価指標による定量的測定が行われており、目標値に対して実績値が提示されているにもかかわらず、それを最後に総合的な観点から主観的に評価する、あるいは2つの部門長の相対的業績を評価対象として主観的に評価する（あるいは一方の業績のみを評価する場合でももう一方の業績を参照できるようにする）意義・必要性については、ほとんど論じられていない。また、実務において共通指標バイアスが生じている事例はいくつか報告されているけれども（Ittner et al., 2003a; Kraus and Lind, 2010）、そもそもなぜそうした主観的評価が行われるのか、その動機については明らかにされていない。

因果連鎖情報の提供やBSCによる多面的な測定が主観的評価の必要性を高めることに繋がると安易に想定することは難しいように思われる。また、標準化された共通指標は、一方の測定結果がもう一方の測定誤差に関する情報を提供するという点で、インフォーマティブであるので（Arya et al., 2005）、共通指標バイアスは問題だけを引き起こしているとは必ずしも言えないかもしれない。

それゆえ、主観的業績評価の利用に大きな役割期待が求められるような状況を想定した実験を行うことが必要になってくるだろう。主観的業績評価の利用は、業績評価指標の不完全さ、管理不能性、操作可能性、環境の予測不可能性によって促進されることが知られている（北田, 2015; Baiman and Rajan, 1995; Bol, 2008）。評価バイアスを扱った研究ではないが、Cheng and Coyte (2014)では、必ずしも業績測定対象とすることが容易ではない行動（知識共有や役割外行動）に対する意欲を高めるインセンティブ・スキームとしての役割期待ができる評価として、主観的業績評価（重みづけ）を取り上げている。そして、主観的な重みづけと戦略マップの活用によって、直接的かつ具体的な形で評価対象としなくても、役割外行動を従業員に動機づけられることを明らかにして



いる。

BSCを用いた業績評価であっても、不完全な業績測定は回避できず、戦略の進捗状況を完全に測定することは難しいとして、主観的に総合評価をする必要があるとも考えることはできる。ただし、この捉え方はBSCの基本思想とは逆行する<sup>11</sup>ので、評価者がBSCを利用する際に主観的な評価という方法を選択する傾向があるのかどうかということ自体が検証対象になりうる。複数の指標による定量的な測定結果を解釈し、総合的観点から主観的に評価値を与えるには、かなりの認知負荷が必要となるので、そもそも先行研究の実験で想定されるような主観的評価自体が回避される傾向があるかもしれないからである。さらに、主観的評価を被験者に強制する実験であっても、主観的評価の目的<sup>12</sup>を明示的に実験設定に反映し、より実際のBSCでの運用状況を睨んだ研究が必要であると思われる。

#### 4.2. 因果連鎖情報の多様性

因果連鎖情報の提供を扱った先行研究の多くが、Banker et al. (2004)の実験設定を参照しており、戦略マップを用いた情報提示の影響を検証してきた。これらマップは、Kaplan and Norton (2001, 2004)の説明と同様に、戦略目標（あるいは視点）間の関係を矢印で結んだものとなっている。ただし、マップは、目標や視点ではなく、業績評価指標間を矢印で結びつけることで描くこともできる。

しかしながら、情報提示という観点からみた、戦略目標と業績評価指標の違いについては、先行研究では検証されておらず、両者を同列に扱って良いのか

---

11 BSCによる戦略とリンクした複数指標の評価は、伝統的な財務指標のみに依存した評価からの脱却を意図しており、測定の不完全さという点で大きな改善が見込めるものである。

12 たとえば、戦略を台無しにしてしまう不測事態による管理不能要因によるノイズの調整やそれら不確実性に関する情報提供の促進は重要な主観的評価の目的の1つである (Simons, 1995)。一部の先行研究 (Cheng and Humphreys, 2012) では、外部情報と戦略との関連性やそれら外部情報に基づいた戦略の適切性といった主観的な評価が重要な役割を果たす場面が注目されている。

は不明瞭のままである。戦略目標とその達成度を測定する業績評価指標を比較すれば、戦略目標に基づいて業績評価指標が設定されるので、評価者の解釈の余地がより大きいのは戦略目標の方であろう。そもそも、戦略目標に対して、それ以上の数の業績評価指標が設定されることも少なくないので、もし指標を用いてマップを描けば、戦略目標で描かれたマップとは異なる構造となる。Carmona et al. (2011) は、戦略マップの構造がBSCを用いた投資意思決定の判断に影響することを明らかにしている<sup>13</sup>。こうした発見も踏まえれば、因果連鎖を何で描写するかによって、評価判断に違いをもたらすのかは検討する価値があると言えよう。

なお、評価バイアスの研究に限定しないのであれば、さらに多様な情報内容の違いが取り扱われている。たとえば、タイムラグ情報の記載 (Humphreys et al., 2016)、因果的な言語の利用 (Loftus and Tanlu, 2018)、因果モデルの正確さ (Kelly, 2010; Huo et al., 2021; Vera-Muñoz et al., 2007) が、意思決定やそれを通じた業績改善、外部情報の評価、従業員の感情に及ぼす影響が検証されている。因果連鎖に関する情報といっても、その内容あるいは表示方法は多様であり、どのような形が望ましいのかは、実践的にも重要な課題であり、拡張の余地があろう。

#### 4.3. 中長期的な時間軸

因果連鎖を示す戦略マップの利用によって、評価者がインナーの視点に関連した指標よりもアウトターの視点に関連した指標をより重視した評価をしてしまうという発見 (Rotaru et al., 2020) は、他の研究では肯定的に評価されている

---

13 Carmona et al. (2011) で注目された戦略マップの構造は、マップ上での4つの視点の結びつき方であり、ピラミッド型とサイロ型の2つに区分されている。前者は、財務の視点を最上位として、マーケティング、環境、イノベーションの視点が順に結びついた、一般的に例示されることの多いマップである。後者は、縦に4つの視点が直結するのではなく、2つの列で並行的に配置されるマップである。実験の結果、ピラミッド型の戦略マップを与えられた場合、サイロ型の戦略マップを与えられた場合よりも、短期的な財務成果を約束する投資案をより好む傾向があることが明らかにされている。

戦略マップの有用性に疑問を投げかけるものである。そのため、彼らの発見がより現実的な状況を想定しても観察されるのか、研究蓄積を図っていく必要があるだろう。なかでも重要と考えられるのが、Rotaruら自身が指摘しているように、より長期的な時間軸を考慮に入れることであろう。

Rotaru et al. (2020) の実験に限られたことではないけれども、1期分のBSCによる業績測定結果に基づいた一度限りの業績評価の判断が考察対象とされることが多い。しかしながら、BSCでは、戦略が完遂したときの最終目標値のみならず、そのマイルストーン目標値も設定される (Kaplan and Norton, 1996, 2001)。そのため、単年度の実績値・目標値のみでは、BSCが実行されている文脈を十分に反映できているとは言えないだろう。評価時点が戦略全体の計画期間のどこに位置しているのか、あるいは最終的な目標水準と現行の目標水準がどの程度乖離しているのかといった情報は、戦略目標の因果性に関する意識や業績評価の判断に影響を与える可能性がある。たとえば、複数の業績評価指標を用いた評価を行うとしても、財務の視点は最終的な目標として位置づけられているので、戦略の計画期間終了時には、他の3つの視点よりも相対的に重要視される可能性が高い (Krumwiede et al., 2013)<sup>14</sup>。特に、下位の視点から上位の視点へと戦略目標がタイムラグを伴い繋がった因果連鎖が組織内で明確に想定されているのであれば、戦略計画の序盤では下位の視点を、終盤では上位の視点を相対的に重視する傾向が強まるかもしれない。Rotaru et al. (2020) では、財務の視点のみが重視されているわけではなく、学習と成長の視点も含めたアウトターの視点を相対的に重要視した評価をすることを追加の分析で確認しているけれども、上述したような状況でも、同様の発見が得られるのかは更なる検証が必要であろう。

---

14 共通指標バイアスへの複数期間にわたる成果のフィードバックが及ぼす影響を検証した Krumwiede et al. (2013) の実験では、1年目の評価時よりも4年目の評価時にて、財務指標を含む共通指標への依存度が増していた。

#### 4.4. 評価対象の違い

管理会計のコンベンショナル・ウィズダムの1つとして、次に示す評価対象に関するものがある。すなわち、資源配分意思決定を目的とする組織の評価では組織に対する追跡可能性を、動機づけを目的とした人の評価では、個人にとっての管理可能性を重要視して、評価システムを構築しなければならない（岡本，1990；小林，1979）。しかしながら、BSC 研究では、これらの区分や違いに対して、ほとんど研究関心が向けられていない。

先行研究の実験では、組織の業績がBSCを用いて測定されており、それに基づき、その組織の長である管理者<sup>15</sup>の業績を評価することが求められている。このとき、当該管理者がBSCで測定される業績に対して、どの程度の管理可能性を有しているのかは明確にされていない。これは、「BSCで実行が意図される事業戦略の策定と実行に包括的な責任を負っている」と考えるのが妥当と思われる、ポジションに就いている管理者（事業部長やユニット長）の業績評価が、分析対象になっているためかもしれない。ただし、Johnson et al. (2014)では、管理者の就任時期と戦略の実行期間との関係から管理可能性を捉えた実験を行っている（注10参照）。3年や5年といった中長期の時間を総合すれば、BSCの全視点のすべての指標が管理可能であり、その実行責任を負うと言えるが、評価のタイミングによっては、遅行する業績評価指標に対する管理可能性を求めるのは妥当ではないかもしれない。つまり、1期で戦略実行が完結することは稀であり、Johnsonらの設定した状況は妥当性を有していると言えよう。なお、時間軸に基づいた管理可能性概念の検討は、上述した中長期的な時間軸の議論とも密接に関わっており、複合的に検討していく必要がある。

管理可能性原則の遵守・逸脱に関しては、創造性の高い問題解決や協力的な行動形成といったポジティブな効果も含めてさまざまな角度から検討が行われ

---

15 BSCにおいては、指標や戦略的イニシアティブを下位組織へと展開するカスケードの議論（Kaplan and Norton, 2001）があり、ここでは最終的に個人レベルにまでの展開が想定されている。しかしながら、先行研究では、こうした下位展開やより現場に近い組織成員の評価は検討対象とされておらず、検討価値のある領域ではないと思われる。

ている (Burkert et al., 2017; Bushman et al., 1995; Giraud et al., 2008; Frow et al., 2005; Merchant, 1989)。これらの効果と評価の公平性をバランスさせることは、BSCによる業績評価においても重要であろう (Cheng and Coyte, 2014)。

また、いくつかの先行研究では、人と組織の評価結果の違いや相互の影響について検討する価値があることを示唆する研究結果を報告している。たとえば、個人業績の評価経験と成果のフィードバックの提供が共通指標バイアスを緩和することを発見した Krumwiede et al. (2013) では、個人ではなく組織業績の評価経験を独立変数とした場合には有意な影響が得られなかったことが確認されている。また、Wong-On-Wing et al. (2007) は、評価者の認知的な限界によって、戦略の影響を考えさせる仕組みがなければ、たとえ戦略の質が低く、業績が低迷している部分があったとしても、それを事業部長（の能力）に原因があると評価する傾向があることを実験によって明らかにした。一方で、戦略の有効性が低いと考える評価者は、部下の業績を評価する際に、戦略に沿った評価基準をあまり考慮しない傾向があることも実証されている (Johnson et al., 2014)。これらの結果は、BSCのようなフォーマットを用いた評価でも、個人と戦略（組織）というように評価対象が異なれば、評価者は異なるメカニズムのもと評価判断を下している可能性があることを示唆している。BSCによる業績評価研究において、管理者と戦略（組織）という評価対象の違いが明示的に独立変数として扱われていないので、まずは対象の違いが評価バイアスと関連しているのかどうか明らかにすることが基礎的な知見として求められるだろう。

## 5. おわりに

本研究ノートでは、BSCに代表される業績評価システムの因果連鎖概念について整理した上で、主要研究潮流の1つである評価バイアスとの関係について検討している先行研究をサーベイした。既存研究の発見によれば、BSCにおける複雑な情報認知・処理の負担を回避しようと、偏った評価行動をとる傾向が認められる一方で、戦略マップによる因果連鎖に関する情報提供はそうし

たバイアスを是正する可能性を有している。そして、提供される情報内容、情報の提示方法の細かな差異によって、評価行動への影響が異なっていることも明らかにされている。本研究ノートでサーベイした文献で検討されていた問題は、認知に関連したバイアスであるが故に、インセンティブ・システムの設計といった動機づけの観点からは解決が期待しにくい。そのため、先行研究によって、評価バイアスの存在に警鐘が鳴らされ、そのための有効な対処方法が提案されてきたことは、理論的のみならず実践的な貢献も高いものと評価することができる。その一方で、前述したように、より現実的な状況を想定する形で研究の拡張余地は残されているように思われる。

なお、因果連鎖情報の提供による意思決定への影響を検証した研究については、紙幅の関係上取り扱うことができなかった。戦略の実行や修正に関連した意思決定は、戦略の評価と深く関連しており、BSCに関する知見を深めていくためには、外部あるいはリスク情報の関連性・重要性の評価（Cheng and Humphreys, 2012; Cheng et al., 2018）や戦略実行に関連した資源配分意思決定（Humphreys et al., 2016）も含めて、一連の評価・判断プロセスの中で検討を試みる必要があるだろう。たとえば、Dalla Via et al. (2019) は、因果連鎖の表示のみならず、どのような会計責任が求められるのかという被評価者の置かれた状況も検討対象として、因果連鎖の提供が、成果型の会計責任のもとでは、情報探索努力を促進し高い意思決定の質を実現するのに適しているが、プロセス会計責任のもとでは、不要な情報探索を削減する効果はあるけれども、意思決定の質を改善するほどではないことを明らかにしている。こうした評価と意思決定の両面を同時に扱うような実験は、BSCの利用により近似した状況での検討が可能となっていると言えるだろう。それゆえ、今後の文献サーベイでは、評価バイアスの議論のみ切り取ることなく、より包括的な観点から知見の取りまとめを行うことが重要となるであろう。

#### 付記

本研究はJSPS 科研費 19K01995 の助成を受けたものです。

## 参考文献

- 伊藤公一朗. 2017. 『データ分析の力：因果関係に迫る思考法』 光文社.
- 岡本清. 1990. 『原価計算』 国元書房.
- 北田智久. 2015. 「主観的業績評価研究に関する現状と課題：文献レビューに基づく考察」 『原価計算研究』 39(1): 54-64.
- 小林哲夫「事業部制組織における業績評価」 溝口一雄責任編集『体系近代会计学Ⅳ 業績評価会計』 中央経済社, 1979, 71-79.
- 西居豪. 2011. 「バランスト・スコアカード研究の系譜と展望」 『会計学研究所報』 25: 1-116.
- 安井翔太. 2019. 『効果検証入門：正しい比較のための因果推論／計量経済学の基礎』 技術評論社.
- Arya, A., Glover, J., Mittendorf, B., and Ye, L. 2005. On the use of customized versus standardized performance measures. *Journal of Management Accounting Research* 17(1): 7-21.
- Baiman, S. and Rajan, M.V. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review* 70: 557-579.
- Banker, R. D., Chang, H., and Pizzini, M. J. 2004. The balanced scorecard: Judgmental effects of performance measures linked to strategy. *The Accounting Review* 79(1): 1-12.
- Banker, R. D., Chang, H., and Pizzini, M. J. 2011. The judgmental effects of strategy maps in balanced scorecard performance evaluations. *International Journal of Accounting Information Systems* 12(4): 259-279.
- Barnabè, F. 2011. A "system dynamics-based Balanced Scorecard" to support strategic decision making: Insights from a case study. *International Journal of Productivity and Performance Management* 60(5), 446-473.
- Bisbe, J., and Otley, D. T. 2004. The effects of the interactive use of management control systems on product innovation. *Accounting, Organizations and Society* 29(8): 709-737.
- Bol, J. C. 2008. Subjectivity in compensation contracting. *Journal of Accounting Literature* 27(1): 1-24.
- Bukh, P. N., and Malmi, T. 2005. Re-examining the cause-and-effect principle of the balanced scorecard. In Jönsson, S. and Mouritsen, J. (eds.) *Accounting in Scandinavia-The Northern Lights*. Malmö, Sweden: Copenhagen Business School Press, 87-113.
- Burkert, M., Fischer, F. M., Hoos, F., and Schuhmacher, K. 2017. The relationship between lack of controllability and proactive work behavior: an empirical analysis of competing theoretical explanations. *Accounting and Business Research* 47(2): 144-171.
- Bushman, R. M., Indjejikian, R. J., and Smith, A. 1995. Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies. *Journal of Accounting Research* 33: 101-128.
- Carmona, S., Iyer, G., and Reckers, P. M. J. 2011. The impact of strategy communications, incentives and national culture on balanced scorecard implementation. *Advances in*



- Accounting* 27(1): 62–74.
- Cheng, M. M., and Coyte, R. 2014. The effects of incentive subjectivity and strategy communication on knowledge-sharing and extra-role behaviours. *Management Accounting Research* 25(2): 119–130.
- Cheng, M. M., and Humphreys, K. A. 2012. The differential improvement effects of the strategy map and scorecard perspectives on managers' strategic judgments. *The Accounting Review* 87(3): 899–924.
- Cheng, M. M., Humphreys, K. A., and Zhang, Y. Y. 2018. The interplay between strategic risk profiles and presentation format on managers' strategic judgments using the balanced scorecard. *Accounting, Organizations and Society* 70: 92–105.
- Dalla Via, N., Perego, P., and van Rinsum, M. 2019. How accountability type influences information search processes and decision quality. *Accounting, Organizations and Society* 75: 79–91.
- Dilla, W. N., and Steinbart, P. J. 2005. Relative weighting of common and unique balanced scorecard measures by knowledgeable decision makers. *Behavioral Research in Accounting* 17: 43–53.
- Epstein, M. J., and Westbrook, R. A. 2001. Linking actions to profits in strategic decision making. *MIT Sloan Management Review* 42: 39–49.
- Frow, N., Marginson, D. and Ogden, S. 2005. Encouraging strategic behaviour while maintaining management control: Multi-functional project teams, budgets, and the negotiation of shared accountabilities in contemporary enterprises. *Management Accounting Research* 16(3): 269–292.
- Gilbert, D. T., and Malone, P. S. 1995. The correspondence bias. *Psychological Bulletin* 117(1): 21–38.
- Giraud, F., Langevin, P., and Mendoza, C. 2008. Justice as a rationale for the controllability principle: a study of managers' opinions. *Management Accounting Research* 19: 32–44.
- Huo, K., Kelly, K., and Webb, A. 2021. The beneficial learning effects of combining a hypothesis-testing mindset with a causal model. *The Accounting Review* forthcoming.
- Humphreys, K. A., and Trotman, K. T. 2011. The balanced scorecard: The effect of strategy information on performance evaluation judgments. *Journal of Management Accounting Research* 23: 81–98.
- Humphreys, K. A., Gary, M. S., and Trotman, K. T. 2016. Dynamic decision making using the balanced scorecard. *The Accounting Review* 91(5): 1441–1465.
- Ittner, C. D., and Larcker, D. F. 2003. Coming up short on nonfinancial performance measurement. *Harvard Business Review* 81(11): 88–95.
- Ittner, C. D., Larcker, D. F., and Meyer, M. W. 2003a. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78(3): 725–758.
- Ittner, C. D., Larcker, D. F., and Randall, T. 2003b. Performance implications of strategic performance measurement in financial services firms. *Accounting, Organizations and Society* 28(7): 715–741.
- Johnson, E. N., Reckers, P. M. J., and Bartlett, G. D. 2014. Influences of timeline and

- perceived strategy effectiveness on balanced scorecard performance evaluation judgments. *Journal of Management Accounting Research* 26(1): 165-184.
- Jones, E. E., and Nisbett, R. E. 1971. *The Actor and the Observer: Divergent Perceptions of the Causes of Behavior*. Morristown, NJ: General Learning Press.
- Kaplan, R. S., and Norton, D. P. 1992. The balanced scorecard: Measures that drive performance. *Harvard Business Review* 70(1): 71-79.
- Kaplan, R. S., and Norton, D. P. 1993. Putting the balanced scorecard to work. *Harvard Business Review* 74(5):134-147.
- Kaplan, R. S., and Norton, D. P. 1996. *The Balanced Scorecard: Translating Strategy into Action*. Boston, MA: Harvard Business School Press.
- Kaplan, R. S., and Norton, D. P. 2001. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Press.
- Kaplan, R. S., and Norton, D. P. 2004. *Strategy Maps: Converting Intangible Assets into Tangible Outcomes*. Boston, MA: Harvard Business School Press.
- Kaplan, R. S., and Norton, D. P. 2006. *Alignment: Using the Balanced Scorecard to Create Corporate Synergies*. Boston, MA: Harvard Business School Press.
- Kaplan, R. S., and Norton, D. P. 2008. *The Execution Premium, Linking Strategy to Operations for Corporate Advantage*. Boston, MA: Harvard Business School Press.
- Kelly, K. 2010. Accuracy of relative weights on multiple leading performance measures: Effects on managerial performance and knowledge. *Contemporary Accounting Research* 27(2): 577-608.
- Kober, R., and Northcott, D. 2021. Testing cause-and-effect relationships within a balanced scorecard. *Accounting & Finance*, 61(S1): 1815-1849.
- Kohavi, R., Tang, D. and Xu, Y. 2020. *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*. Cambridge University Press. 大杉直也訳. 2021. 『A/Bテスト実践ガイド 真のデータドリブンへ至る信用できる実験とは』KADOKAWA.
- Kraus, K., and Lind, J. 2010. The impact of the corporate balanced scorecard on corporate control: A research note. *Management Accounting Research* 21(4): 265-277.
- Krumwiede, K. R., Swain, M. R., Thornock, T. A., and Eggett, D. L. 2013. The effects of task outcome feedback and broad domain evaluation experience on the use of unique scorecard measures. *Advances in Accounting* 29: 205-217.
- Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108(3): 480-498.
- Leigh, A. 2018. *Randomistas: How Radical Researchers are Changing our World*. Yale University Press. 上原裕美子監訳. 2020. 『RCT 大全：ランダム化比較試験は世界をどう変えたのか』みすず書房.
- Libby, T., Salterio S. E., and Webb, A. 2004. The balanced scorecard: The effects of assurance and process accountability on managerial judgement. *The Accounting Review* 79(4): 1075-1094.
- Lipe, M. G., and Salterio, S. E. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75(3): 283-298.
- Loftus, S., and Tanlu, L. J. 2018. Because of "because": Examining the use of causal language in relative performance feedback. *The Accounting Review* 93(2): 277-297.

- Malina, M. A., Nørreklit, H., and Selto, F. H. 2007. Relations among measures, climate of control and performance measurement models. *Contemporary Accounting Research* 24(3): 935–982.
- Malina, M. A., and Selto, F. H. 2001. Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard. *Journal of Management Accounting Research* 13(1): 47–90.
- Merchant, K. A. 1989. *Rewarding Results: Motivating Profit Center Managers*. Cambridge, MA: Harvard Business School Press.
- Nørreklit, H. 2000. The balance on the balanced scorecard: A critical analysis of some of its assumptions. *Management Accounting Research* 11(1): 65–88.
- Nørreklit, H. 2003. The balanced scorecard: What is the score? A rhetorical analysis of the balanced scorecard. *Accounting, Organizations and Society* 28: 591–619.
- Rehder, B. 2014. Independence and dependence in human causal reasoning. *Cognitive Psychology* 72: 54–107.
- Roberts, M. L., Albright, T. L., and Hibbets, A. R. 2004. Debiasing balanced scorecard evaluations. *Behavioral Research in Accounting* 16(1): 75–88.
- Rotaru, K., Fehrenbacher, D. D., Liang, M. H., and Schulz, A. K. D. 2020. Causal inference in judgment using the balanced scorecard. *Journal of Management Accounting Research* 32(2): 201–224.
- Simons, R. 1995. *Levers of Control: How Managers Use Innovative Control Systems to Drive Strategic Renewal*. Boston: Harvard Business School Press.
- Speckbacher, G., Bischof, J., and Pfeiffer, T. 2003. A descriptive analysis on the implementation of Balanced Scorecards in German-speaking countries. *Management Accounting Research* 14(4): 361–387.
- Speklé, R. F., van Elten, H. J., and Widener, S. K. 2017. Creativity and Control: A Paradox—Evidence from the Levers of Control Framework. *Behavioral Research in Accounting* 29(2): 73–96.
- Sundin, H., Granlund, M., and Brown, D. A. 2010. Balancing multiple competing objectives with a balanced scorecard. *European Accounting Review* 19(2): 203–246.
- Taylor, W. B. 2010. The balanced scorecard as a strategy-evaluation tool: The effects of implementation involvement and a causal-chain focus. *The Accounting Review* 85(3): 1095–1117.
- Vera-Muñoz, S., Shackell, M., and Buehner, M. 2007. Accountants' usage of causal business models in the presence of benchmark data: A note. *Contemporary Accounting Research* 24(3): 1015–1038.
- Webb, R. A. 2004. Managers' commitment of the goals contained in a strategic performance measurement system. *Contemporary Accounting Research* 21(4): 925–958.
- Wong-On-Wing, B., Guo, L., Li, W., and Yang, D. 2007. Reducing conflict in balanced scorecard evaluations. *Accounting, Organizations and Society*, 32(4–5): 363–377.
- Yopchick, J. E., and Kim, A. S. 2009. The influence of causal information on judgments of treatment efficacy. *Memory & Cognition* 37(1): 29–41.