

評価学博士

Ryo SASAKI, Ph. D.

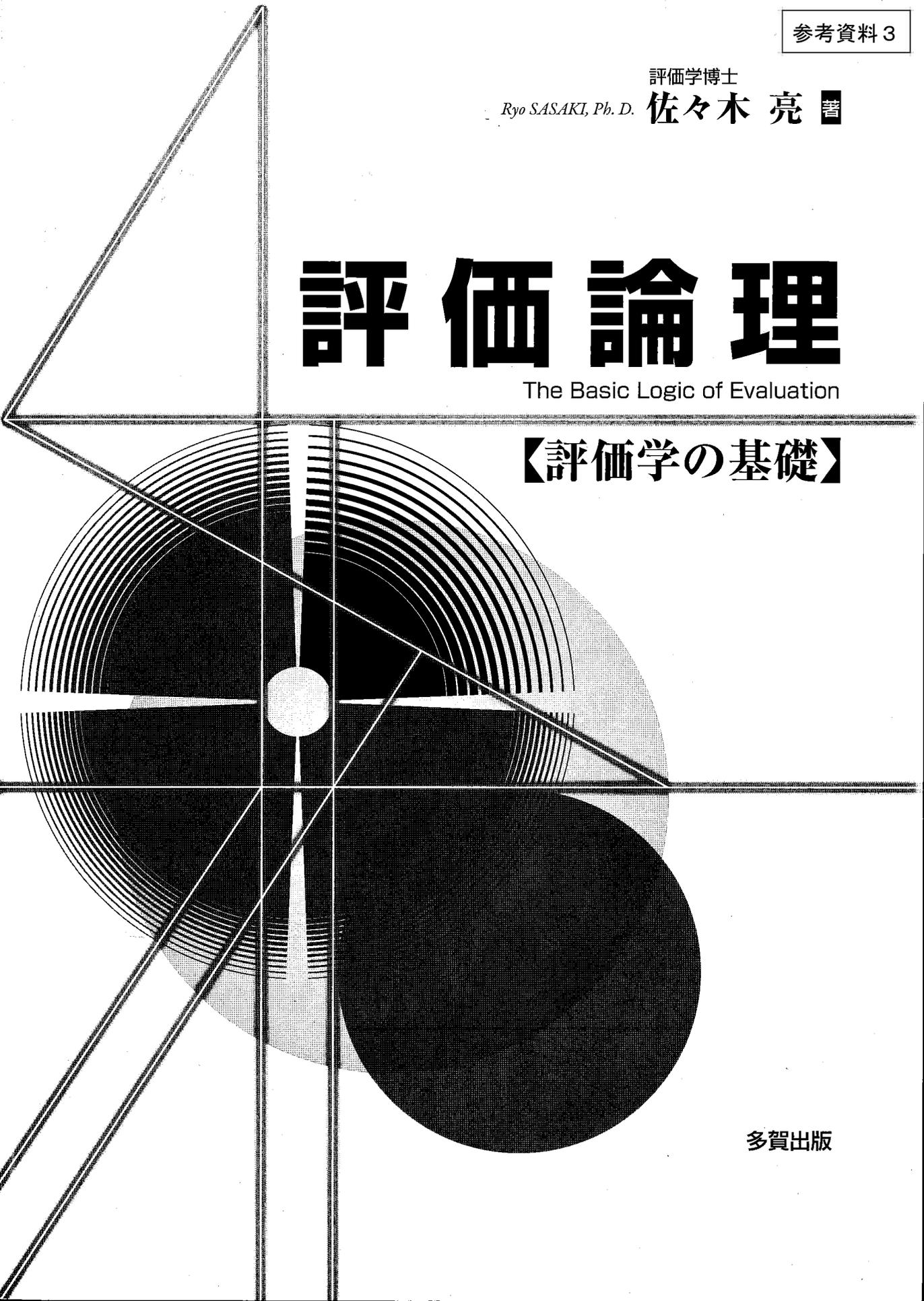
佐々木 亮 著

評価論理

The Basic Logic of Evaluation

【評価学の基礎】

多賀出版



2. 因果関係を特定するための調査デザイン（定量的手法）

因果関係を特定する方法としては、定性的方法と定量的方法がある。定量的方法としては、もっとも簡易な方法と言われる専門的判断および事前・事後比較デザインから、もっとも客観的と言われる実験デザインまで整理されている。以下の表にその要約を掲載した。それぞれのデザインの利点、制約、利用上の留意点に関する詳細に興味のある方は、出典をあたらせたい（龍・佐々木（2004）『第5章インパクト評価』）。

因果関係性の特定のための評価デザインの類型その1（定量的手法）

評価デザイン名 / 概念図	信頼性
<p>① 専門家判断（Expert Judgment）</p> <p>特定分野の知識と経験を有する専門家が、自身の経験に基づいて設定した心の中の基準と実際の状況を比べてインパクトを推定する。例：「私の経験を踏まえると、この村では約2倍の生産性が実現していると思われます。」</p> <p>因果関係の特定の方法としてはまったく勧められないが、広く使われている。</p>	低
<p>② 事前・事後比較デザイン（Before-After Design）</p> <p>シンプルに、事前、事後の指標値を比較し、差があれば因果関係があったと推定する。簡便なので広く用いられている。ただし、事前・事後の間に発生した外部要因による影響値をまったく取り除けないので、因果関係の推定の信頼性は低い。</p> <div style="text-align: center;"> <h3>シンプル事前-事後モデルの概念図</h3> </div> <p style="text-align: right;">出所)筆者作成</p>	↓

因果関係性の特定のための評価デザインの類型その1 (定量的手法)

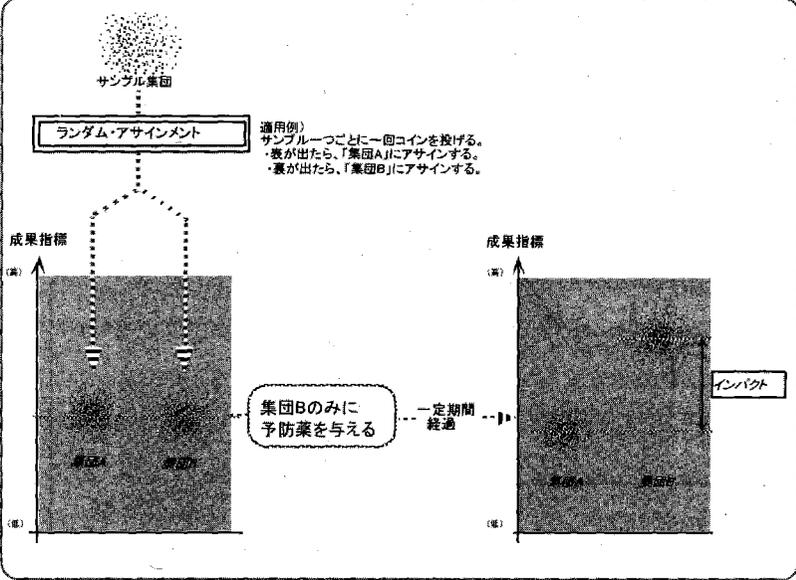
評価デザイン名 / 概念図	信頼性
<p>③ 時系列デザイン (Interrupted Time-Series)</p> <p>施策介入前の長期的トレンドを導き出し、施策介入後にトレンドが変わっていれば、因果関係の存在を推定する。ただし、長期的トレンド以外の外部要因による影響値を取り除けないので、信頼性はそれほど高くない。</p> <p style="text-align: center;">時系列モデルの概念図</p> <p style="text-align: right;">(出所) 筆者作成</p>	中
<p>④ クロスセクションデザイン (Cross Section)</p> <p>地域間の施策実施度合のばらつきとアウトカムのばらつきの相関関係を推定することにより、因果関係の存在を推定する。ある程度の信頼性を確保できる。</p>	↓

(出所) Rossi, Freeman and Lipsey; *Evaluation: A Systematic Approach 6th ed*, p261 の表を参照して筆者作成。

（定量的手法（続））

評価デザイン名 / 概念図 / 解説	信頼性
<p>⑤ 一般指標デザイン (Generic Control)</p> <p>全国平均値、全県平均値などの一般指標値を比較に用いる。外部要因による影響値をある程度除去して考えることができるので（なぜなら対象地域が受けた影響とある程度同じ影響を一般指標値も受けているはずだから）、因果関係の存在の特定に関してある程度の信頼性を確保できる。わりと簡単に用いることができる。</p> <p style="text-align: center;">一般指標モデルの概念図</p>	<p>やや高</p>
<p>⑥ マッチング・デザイン (Matched Control)</p> <p>可能な限り近似のグループを選定して比較に用いる。外部要因による影響はどちらのグループも同程度に受けると考えられるので、因果関係の存在の特定のために高い信頼性を確保できる。</p> <p style="text-align: center;">マッチング・モデルの概念図</p>	<p>高</p>

(定量的手法 (続))

評価デザイン名 / 概念図 / 解説	信頼性
<p>⑦ 実験デザイン (Experimental Design)</p> <p>施策の実施前に、政策適用を無作為割付 (ランダム・アサインメント) により、実施グループと比較グループに分ける。因果関係の存在の特定に関してたいへん高い信頼性を誇るとされる。ただし実際の適用は非常に難しい。</p>  <p>ランダム・アサインメント</p> <p>適用例) サンプル一つごとに一回コインを投げる。 ・表が出たら、「集団A」にアサインする。 ・裏が出たら、「集団B」にアサインする。</p> <p>成果指標 (高)</p> <p>成果指標 (高)</p> <p>集団Bのみに 予防薬を与える</p> <p>一定期間 経過</p> <p>インパクト</p>	<p>極めて 高い</p>

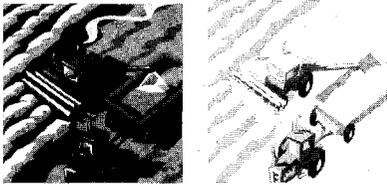
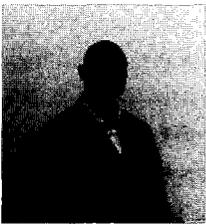
(定量的手法 (補足))

(参考) 数値目標との比較 (目標管理アプローチ)	信頼性
<p>数値目標と比較する。いわゆる業績測定 (Performance Measurement) の応用で、あらかじめ設定した数値目標の達成を成功・失敗の判断基準にする。因果関係の特定とは無縁。</p>	<p>議論の 範囲外</p>

3. 因果関係を特定するための調査デザイン (定性的手法)

すでに述べたとおり、因果関係を特定する方法としては、定性的方法と定量的方法がある。定性的方法としては、直接観察、情報保持者インタビュー、焦点化集団インタビュー、参与観察、アンケート調査 (サンプル、悉皆)、文書分析、行政データ分析がある。また最近の手法として、写真やビデオを利用するという方法もある。それぞれのデザインの利点、制約、利用上の留意点に関する詳細に興味のある方は、出典をあたられたい (Yin, R. (1994). *Case Study Research: Design and Methods*, 2nd edition, など)。

因果関係性の特定のための評価デザインの類型その2（定性的手法）

<p>① 直接観察（Direct observation）</p> <p>評価者が現場訪問して、自分の目で観察して因果関係を判断する。前提条件として、評価者が対象分野の知識と経験を有していなければ成立しない方法である。</p> <p>例1 「灌漑が米の生産増加を実現している」という因果関係を自分の目で見て判断する。</p>  <p>例2 上記の例で、灌漑をしている地域と、していない地域の2箇所を訪問して自分の目で見て、その違いから判断する。</p> 	<p><強み></p> <ul style="list-style-type: none"> ・現時点での事象を扱うことができる（＝過去のデータが必要ない）。 ・関連する状況も同時に観察できる（＝予期せぬことを発見できる可能性大）。 <p><弱み></p> <ul style="list-style-type: none"> ・時間がかかる（＝実際に現場まで足をのばさねばならない）。 ・選択的になる（幅広い事例を含むことはできない）。 ・観察による影響（＝観察されることで人々の行動が改まったりする可能性がある）。
<p>② 情報保持者インタビュー（キー・インフォーマント・インタビュー）（Key Informant Interview）</p> <p>情報を持っている人（キー・インフォーマント）に評価者がインタビューして因果関係を判断する。（情報を持っている人を、状況を継続的に観察してきたモニターとみなしてインタビューするということ）</p> <p>例 「2年前に来た新任の校長先生の取り組みによって学校は変わったと思いますか？」と、長年勤めている用務員さんに聞く。</p> 	<p><強み></p> <ul style="list-style-type: none"> ・常駐のモニターの代わりとして情報を提供してもらうのでコストが安い。 ・洞察に富む。人間によって認知された因果関係が聞き出せる。 <p><弱み></p> <ul style="list-style-type: none"> ・情報保持者と、事業実施者（＝左の例だと校長先生）の人間関係が回答に影響を与える可能性 ・よく思い出せない為に生じる不正確さ ・面接者が聞きたいと思っていることを情報保持者が察して答える可能性（再帰性という）

(定性的手法 (続))

③ 焦点化集団インタビュー (フォーカス・グループ・インタビュー) (Focus Group Interview)

数名の典型的なユーザーあるいは受益者に集まってもらって輪になってもらい、評価者がインタビューする。

例 「①現在の学校の状況はどうでしょうか? ②年前に新任の校長先生が来てから学校は変わったと思いますか? ③それはなぜですか?」というようにインタビューする。このように、現状、過去との比較、原因と3段階で聞くと、聞かれた人も頭の中を整理しながら答えられる。

(情報保持者インタビューではすでに状況に詳しい人に聞くのでいきなり②だけ聞いていることになる)



<強み>

- ・常駐のモニターの代わりとしてユーザー (受益者) に情報を提供してもらうのでコストが安く済む。
- ・答える人がお互いの情報を確認し合いつつ答えることができる。
- ・洞察に富む一人間によって認知された因果関係が開き出せる。

<弱み>

- ・よく知らなかったり思い出せないために生じる不正確さ
- ・典型的と思われたがじつは特殊なユーザー (受益者) が入り込んでインタビューに影響を与える可能性
- ・面接者が聞きたいことをユーザー (受益者) が答える可能性 (再帰性という)

④ 参与観察 (Participatory observation)

評価者が地元の行事 (イベント) に自ら参加して、自分で体験して因果関係を判断する。

例1 農村の「寄り合い」に参加して、農村の結束力が事業の成功につながっているかどうかを判断する。

例2 灌漑地域の農村に一年住み込んで一緒に農作業をして、灌漑が稲作の増収につながっているかどうかを判断する。



<強み>

- ・対人行動とその動機への洞察が可能。
- ・現時点での事象を扱うことができる (=過去のデータが必要ない)。
- ・関連する状況も同時に観察できる (=予期せぬことを発見できる可能性大)。

<弱み>

- ・評価者が参加することによる影響 (=観察されることで人々の行動がいつもと違う行動になる可能性あり)。
- ・時間がかかる (=実際に現場まで足をのばして長期間観察せねばならない)。
- ・選択的になる (幅広い事例を含むことはできない)。

（定性的手法（続））

⑤ アンケート調査：サンプル抽出型（サンプルサーベイ）（Sample Survey）	
<p>一定の方法で抽出したサンプルに対してアンケート用紙を配って答えてもらう。アンケート結果を分析して、因果関係を判断する。</p> <p>例 「子供手当」という政策が、参考書の購入増加や塾通いの増加などにつながったかどうかを、アンケート調査によって調べる。1万世帯のうち、等間隔で500世帯を抽出して調査する。</p>	<p style="text-align: center;">＜強み＞</p> <ul style="list-style-type: none"> ・直接観察などのように選択的にならず、全体像を把握することができる。 ・定量的なデータが収集可能で、データの信頼区間も計算できる（しなくてもいいが）。 <p style="text-align: center;">＜弱み＞</p> <ul style="list-style-type: none"> ・アンケート調査の実施に費用と時間がかかる（評価者の費用、協力する対象機関の費用の両面がある）。 ・アンケート集計結果の分析に費用と時間がかかる。 ・回収率が低かった場合は、集計結果にバイアスがかかっている可能性がある。
⑥ アンケート調査：全数型（悉皆サーベイ、センサスとも言う）（Population Survey, Census）	
<p>対象者全員に対してアンケート用紙を配って答えてもらう。アンケート結果を分析して、因果関係を判断する。</p> <p>例 「子供手当」という政策が、参考書の購入増加や塾通いの増加などにつながったかどうかを、アンケート調査によって調べる。1万世帯のうち、1万世帯の全世帯に対して調査する。</p>	<p style="text-align: center;">＜強み＞</p> <ul style="list-style-type: none"> ・アンケート調査：サンプル抽出型と同一。 <p style="text-align: center;">＜弱み＞</p> <ul style="list-style-type: none"> ・アンケート調査：サンプル抽出型と同一。とくに膨大な費用と時間がかかることに留意。 ・通常は現実的ではない。

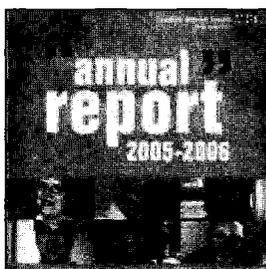


(定性的手法 (続))

⑦ 文書分析 (Documentation Analysis)

対象機関が発行している年次報告書、ニュースレター、会報、公開しているホームページなどを分析して、事業と成果の間に因果関係があるかどうかを判断する。

例 ある学校において、過去10年の年次報告書を時系列的に分析して、教育改革プログラムと成果の因果関係を判断する。



<強み>

- ・安定的—繰り返し戻って分析が可能
- ・もともと評価のためにとして作成されたものではないので中立的。
- ・名称、参考文献、事象の詳細が正確
- ・長期間、多くの事象、多くの状況が広範囲に収集されている可能性大。

<弱み>

- ・文書の執筆者の個性や編集方針により文書の質がばらばらになっている可能性あり。
- ・紙ベースの場合が多く検索可能性が低い。
- ・収集が不完全な場合、分析にバイアスがかかる可能性あり。
- ・公表されていないデータによるバイアスの可能性あり。
- ・文書を提供してもらえない可能性あり。

⑧ 行政データ分析 (Documentation Analysis)

対象機関が保存しているサービス記録、出席者記録、活動記録、アンケート結果などを分析して、事業と成果の間に因果関係があるかどうかを判断する。

例 ある専門学校において、過去2年の出席生徒数の記録、個人の学力テスト記録、通信簿の一覧などを分析して、経営改革プログラムと成果の因果関係を判断する。



<強み>

- ・文書分析 (Documentation Analysis) と同一。

<弱み>

- ・文書分析 (Documentation Analysis) と同一。
- ・プライバシーへの配慮により、データを提供してもらえない可能性が、文書分析の場合よりも大きい。

（定性的手法（続））

⑨ 写真やビデオの利用 (The use of photos and videos)	
<p>同一の場所に関して、施策の事前段階で撮影していた写真と、施策の事後段階で撮影した写真を見比べて、因果関係を判断する。写真の代わりにビデオを利用することもある。</p> <p>例 「ゴミのポイ捨てをやめよう」というキャンペーンに関して、事前の道路の状況の写真と事後の道路の状況の写真を見比べて、キャンペーンと道路の綺麗さの因果関係を判断する。</p> <p style="text-align: center;">(事前) (事後)</p> <div style="display: flex; justify-content: space-around;">   </div>	<p style="text-align: center;"><強み></p> <ul style="list-style-type: none"> 一人の専門家の判断に頼るのではなく、多くの人たちが同時に見比べて判断することが可能。 <p style="text-align: center;"><弱み></p> <ul style="list-style-type: none"> 人間を撮影対象とする場合には、プライバシーへの配慮が必要。撮影される人に事前に承諾を取る必要があるが、承諾が取れない場合がある。

(出所) Yin, R. (1994). *Case Study Research: Design and Methods, 2nd edition*、佐々木 (2010) 立教大学 21 世紀社会デザイン研究科での授業の配布資料

4. 実験デザインを巡る論争³

(1) 実験デザインの基本原理

内的妥当性を検証するためにもっとも望ましいと考えられている調査デザインは、「実験デザイン」(Experimental Design) と呼ばれるものである。実験デザインは、たとえば被験者 100 人をコインの裏と表で 50 人ずつのグループにランダム（無作為）に分けて、片方のグループだけにある介入行為を適用し、一定期間後に、想定した成果指標（Outcome indicators）に関する差を測定するというデザインである。ランダム（無作為）に分けているので、二つのグループの全ての特徴の平均値が現在と将来について一致しているので（健康指標、社会指標、学歴指標、生涯賃金、置かれた社会状況などすべての指標値に関してこれが言える）、二つのグループに現れた成果指標の差は、途中の唯一の違いである介入行為を適用されたか否かによって引き起こされたときのみとみなすことができる（Greenberg & Shroder, 1997, p.xvii）（図表 2）。

³ 本節 (4. (1)-(4)) のこれ以降の初出は、行政管理研究センター「評価クォーターリー」第 10 号（2009 年 7 月）である。