

平成30年度政策評価に関する統一研修
(名古屋会場)

「政策評価の総論」

平成30年12月11日(火)

国際開発センター主任研究員
佐々木 亮

I あらためて「評価」とは

...

Evaluation revisited

評価とは

「評価は物事のメリット、値打ち、意義を体系的に明らかにすることである」 (Scriven, 1991:139)

“Evaluation is the systematic determination of merit, worth and significance of things”

評価 = 事実特定 + 価値判断
(Evaluation) (Factual identification) (Value determination)

出所: 佐々木亮(2015)評価士養成講座テキスト

評価と調査の目的の違い

調査の目的は、因果関係を含む物事の現象を明らかにすることである(=「事実特定」と考えることができる。

一方、明らかにされた事実に関して何らかの「価値判断」を下して初めて、評価足りえると言える。

(出所) Scriven, M. (1980). *The Logic of Evaluation* Inverness. Edgepress. CA.

価値判断をするためには、「価値を表現する言葉」で結論が述べられねばならない。以下に例を挙げる。

価値を表現する言葉の例(等級付け)(例)

■ たいへん満足だった	■ (実施するだけの) たいへんな価値があった	■ たいへん有意義だった
■ 満足だった	■ (実施するだけの) 価値があった	■ 有意義だった
■ どちらとも言えない	■ どちらとも言えない	■ どちらとも言えない
■ 不満足だった	■ (実施するだけの) 価値はなかった	■ 有意義ではなかった
■ はなはだ不満足だった	■ (実施するだけの) 価値は全くなかった	■ はなはだ有意義ではなかった

また、以下のような簡便な表現も認められる。ただし、別途、それぞれの等級が意味することを明確に定義しておくべきである。

簡便な等級付け

優	5	A	★★★
良	4	B	★★
可	3	C	★
不良	2	D	(星なし)
不可	1	E	

価値判断のための基本的な評価枠組の例

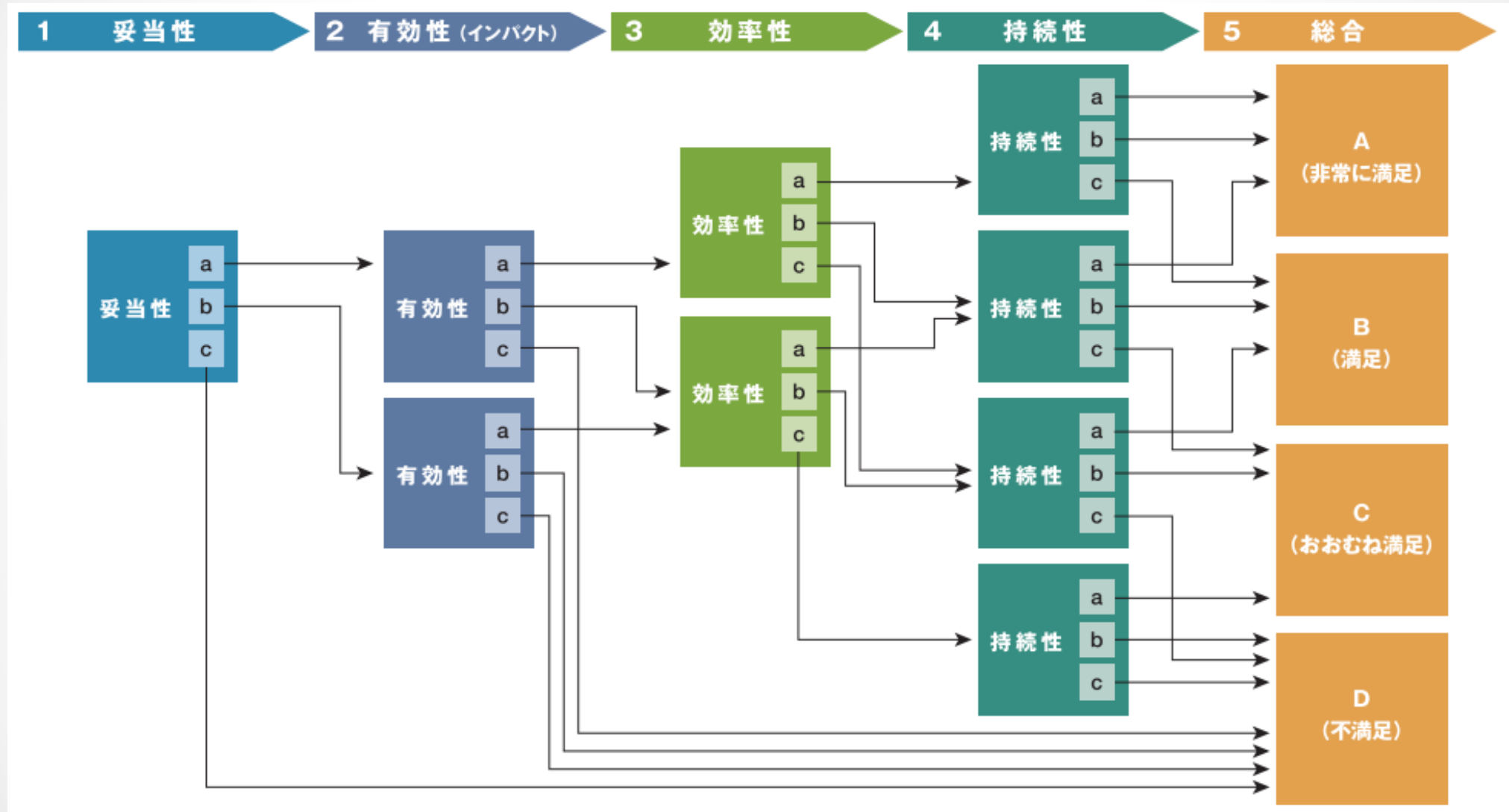
(i) 価値項目 (Criteria of value)

(ii) 価値水準 (Standard of value)

	妥当性	有効性	効率性	持続性
a. 非常に満足	(定義文)	(定義文)	(定義文)	(定義文)
b. 満足	(定義文)	(定義文)	(定義文)	(定義文)
c. どちらとも言えない	(定義文)	(定義文)	(定義文)	(定義文)
d. 不満足	(定義文)	(定義文)	(定義文)	(定義文)
e. 非常に不満足	(定義文)	(定義文)	(定義文)	(定義文)

* 価値項目 & 価値水準を、評価項目 & 評価基準と表現することもある。

総合評価を得るための方法の例(フローチャート)



(出所) 国際協力銀行(2006)『円借款事業評価報告書』

評価の目的(1)

- 形成的評価と総括的評価 (Scriven)
 - ① 形成的評価 (Formative evaluation)
 - 「コックが調理中にスープを味見すること」
 - ・事業や活動の開発・形成
 - ・継続的改善
 - ・立案、実施段階で行う評価
 - ② 総括的評価 (Summative evaluation)
 - 「客が出来上がったスープを味わうこと」
 - ・事業や活動の成果の把握
 - ・プログラム継続・縮小・拡大の判断
 - ・一定期間経た段階で行う評価



相互に補完し合うことで、より効果的なプログラム形成につながる

評価の目的(2)

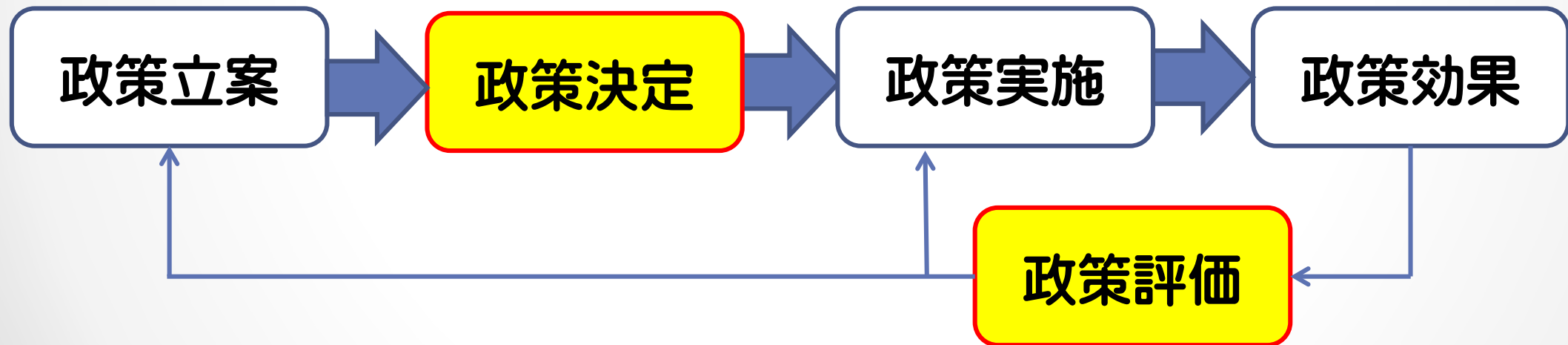
- 評価情報を何のために使うのか (Scriven, 2004)
 - 事業や活動の改善を目的としたもの(形成的評価)
 - 介入のアカウンタビリティの確保を目的としたもの (総括的評価)
 - 科学的な知識・知見を生み出すことを目的としたもの
- その他にもあるが、重要なことは、評価の「目的」をまず明らかにすること。それに合わせて評価方法(誰が、いつ、どのような手法で行うのか)を選択する必要がある。

政策評価の目的

●「政策評価」とは何か？

○政策評価活動は、「情報提供活動」の一種です。

→「政策評価」は、「政策決定」とは異なります。



- 政策実施の改善へのフィードバック ➡ 形成的評価
- 政策立案へのフィードバック ➡ 総括的評価

Ⅱ プログラム評価について

...

Program Evaluation

プログラム (Program) とは

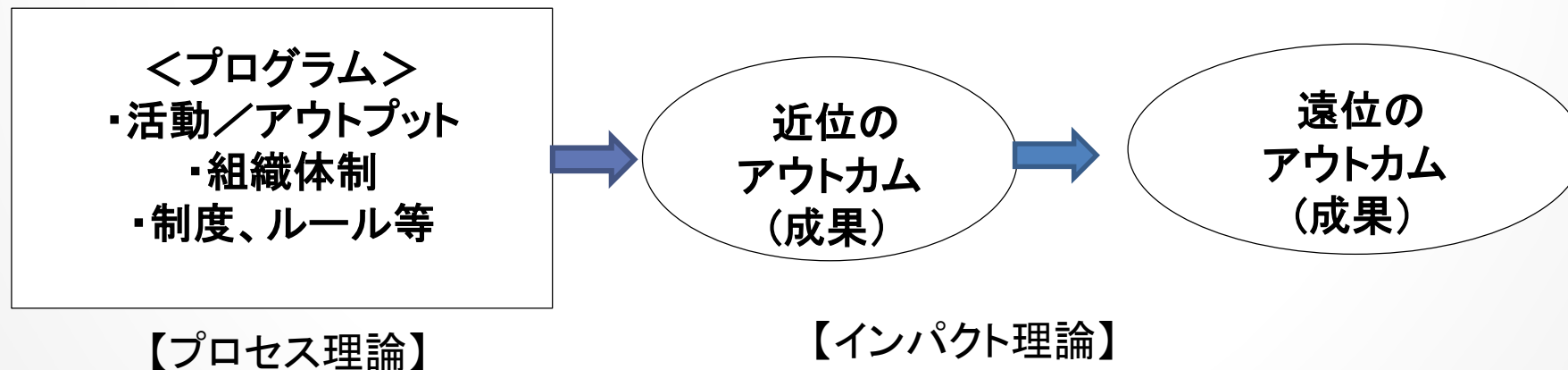
- 「プログラム」とは
 - プログラムとは、ある社会目的を達成するための一連の活動群をさし、それらの活動を実施に導くルール、制度、組織体制や人材を含む取組み全体をさすものである。
 - 特定の組織が主体となって実施する場合もあれば、複数の組織・関係者が協働で実施する取組みもある

プログラム評価の定義

- プログラム評価とは、社会調査の方法を活用し、社会プログラムによる介入の有効性を体系的に調査し、評価を行うこと。その評価は、プログラムを取り巻く政策的・組織的な文脈を考慮した方法で行われるもので、社会状況を改善するための活動の情報源となるものである(Rossi, Lipsey & Freeman, 2004: 16)
- 評価とは、プログラムや政策の改善に寄与するための手段として、明示的または暗示的な基準と比較しながら、プログラムや政策の実施やアウトカムを体系的に査定することである(Weiss, 1998:4)

プログラム理論とは

- プログラム理論 (Program Theory) という考え方
 - ・「プログラムの実施と成果との間に存在する媒介メカニズム」(Weiss)
 - ・プログラムの作戦や活動内容を決定するにあたり活用できる考え方であり、「めざす成果を得るためには、おそらくこういう活動が必要であろう」という「仮説」である
 - ・その「仮説」が有効か、有効にするにはどうしたらよいか、を問う
 - ・「インパクト理論」と「プロセス理論」のふたつの理論が包含される



ロジック・モデル

●ロジック・モデルとは何か？

→プログラム理論を踏まえてプログラムを表現する汎用的枠組み(世界標準)である。

●なぜ、ロジック・モデルが必要なのか？

→プログラムの運営方法や実施プロセスを明らかにすることでブラック・ボックス評価を回避できる。(インパクト理論のみならずプロセス理論の明確化)

→介入のメカニズムを明らかにすることで体系的な評価(5階層)が可能になる。

→目的の設定、指標の検討を行うときの枠組みとして活用できる。

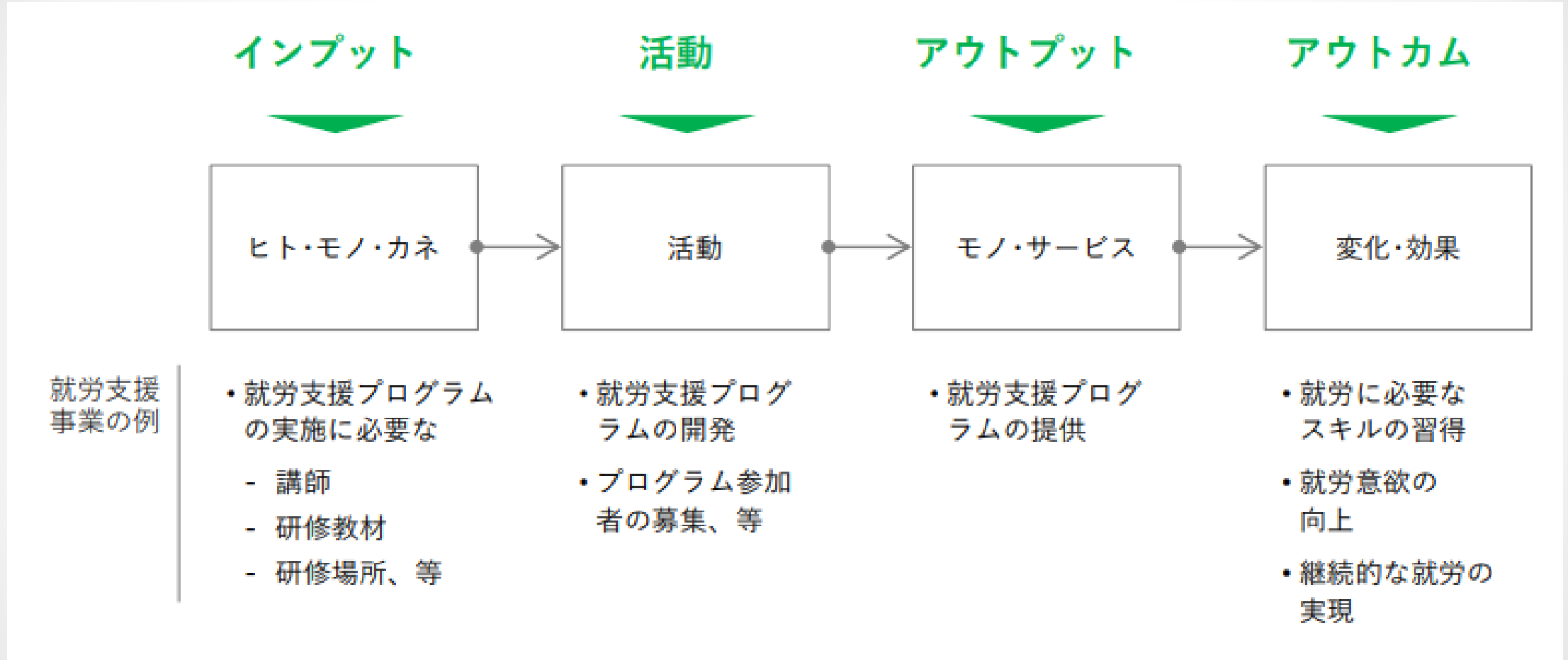
→ステークホルダー間で、プログラムのコンセプト、作戦を検討し、共有するためのツールとして活用できる

●ロジック・モデルの構成要素は何か？

→4つの基本要素でプログラムを表現するものである。

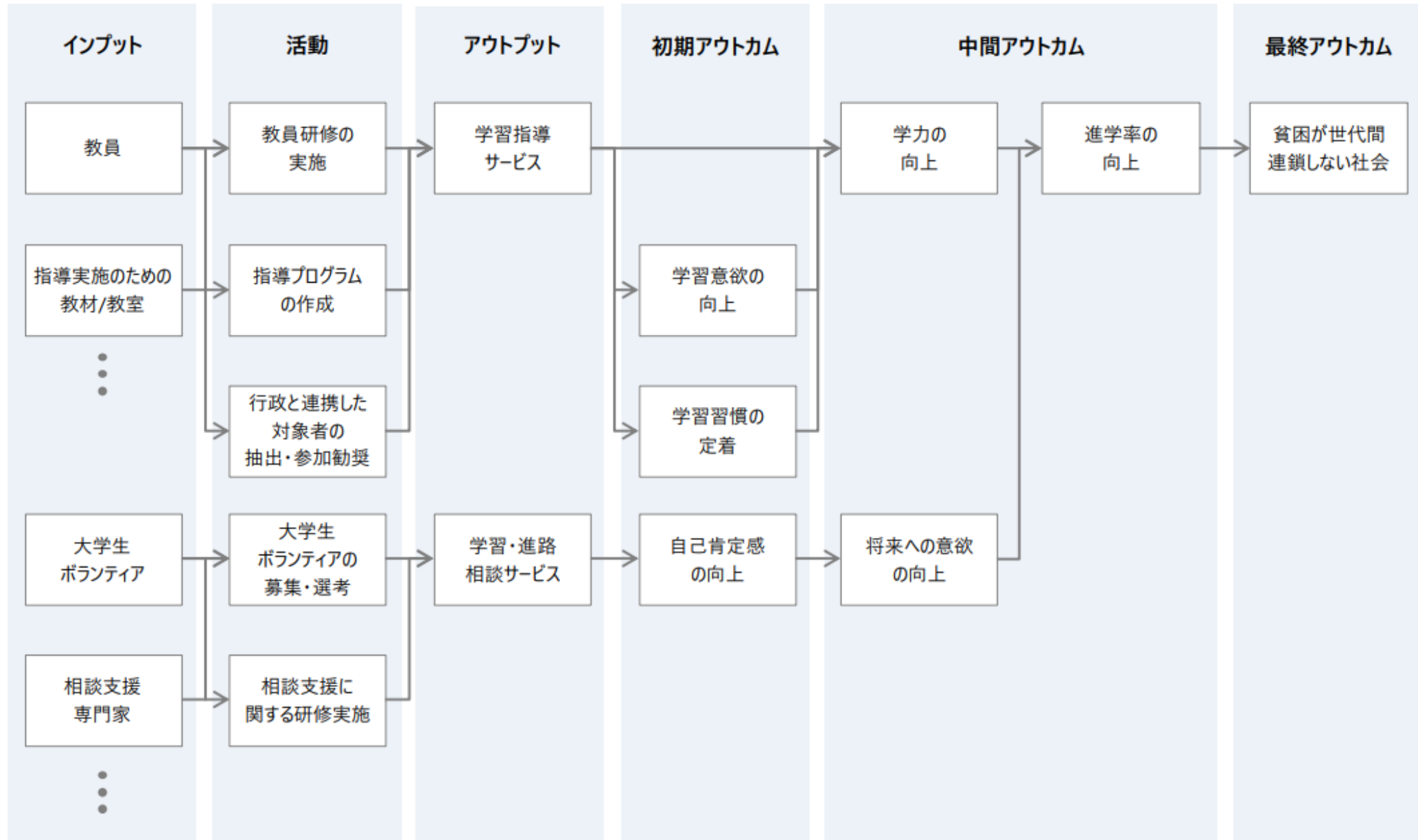
•「インプット」「アクティビティ」「アウトプット」「アウトカム」

就労支援事業のロジックモデルの例



(出所) 日本財団(2017) ロジックモデル作成ガイド

1. 学習支援事業におけるロジックモデルの例



プログラム評価の5階層～体系的に介入を評価する～

(Rossi, Lipsey and Freeman, 2004: 80)

プログラムのコストと効率性の評価
Assessment of Program Cost and Efficiency

プログラムのアウトカム／インパクト評価
Assessment of Program Outcome/Impact

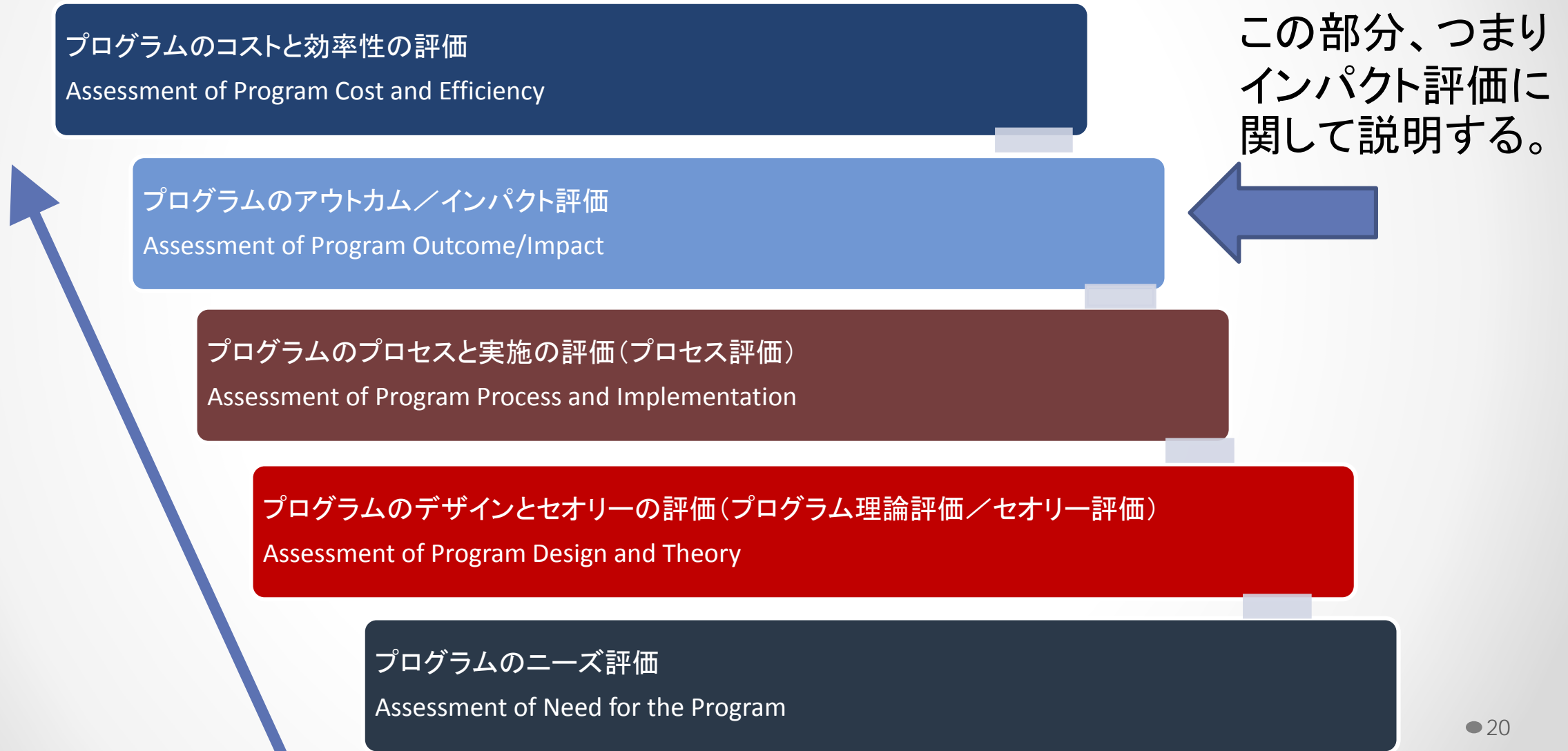
プログラムのプロセスと実施の評価(プロセス評価)
Assessment of Program Process and Implementation

プログラムのデザインとセオリーの評価(プログラム理論評価／セオリー評価)
Assessment of Program Design and Theory

プログラムのニーズ評価
Assessment of Need for the Program

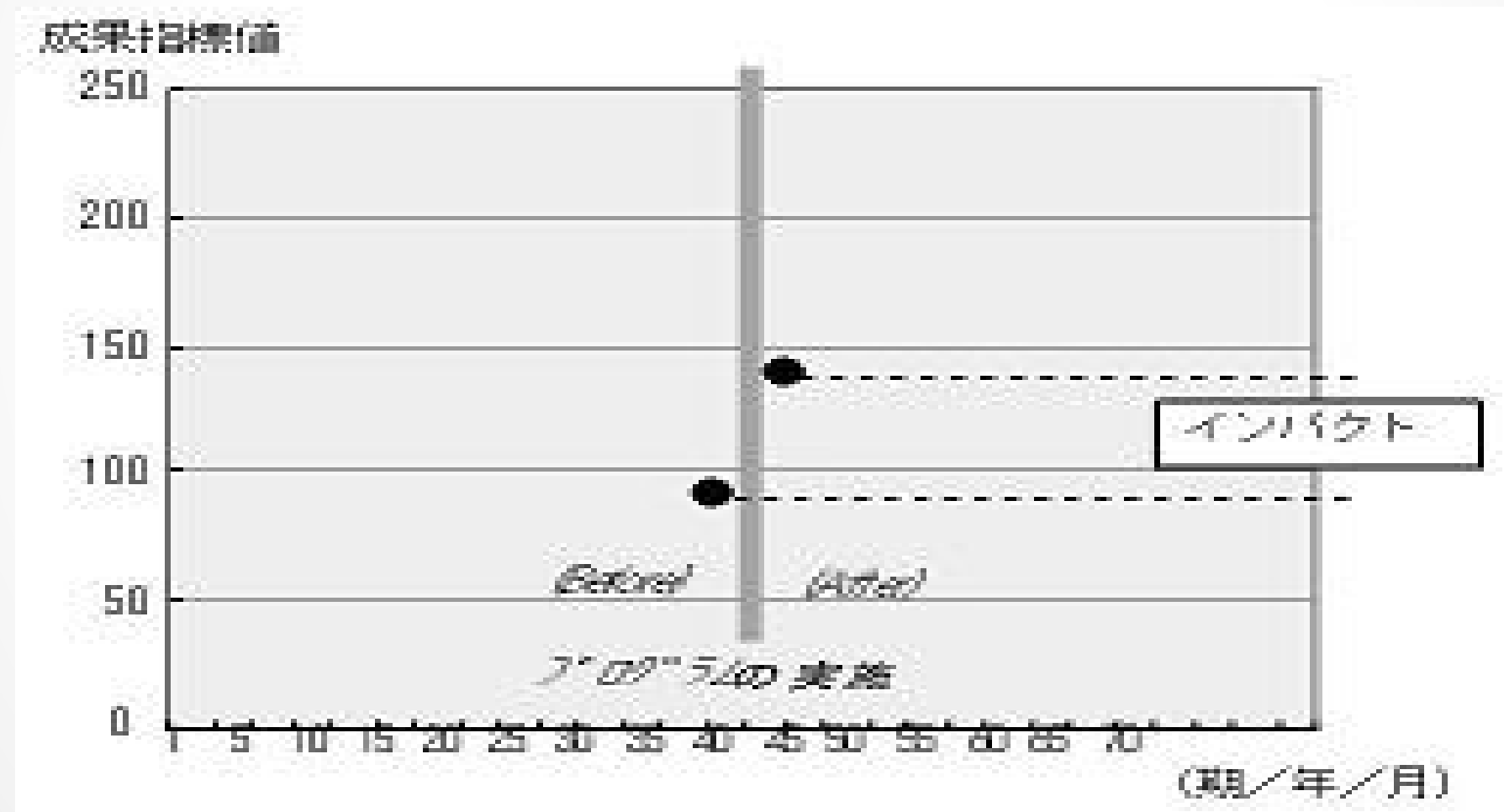
プログラム評価の5階層～体系的に介入を評価する～

(Rossi, Lipsey and Freeman, 2004: 80)



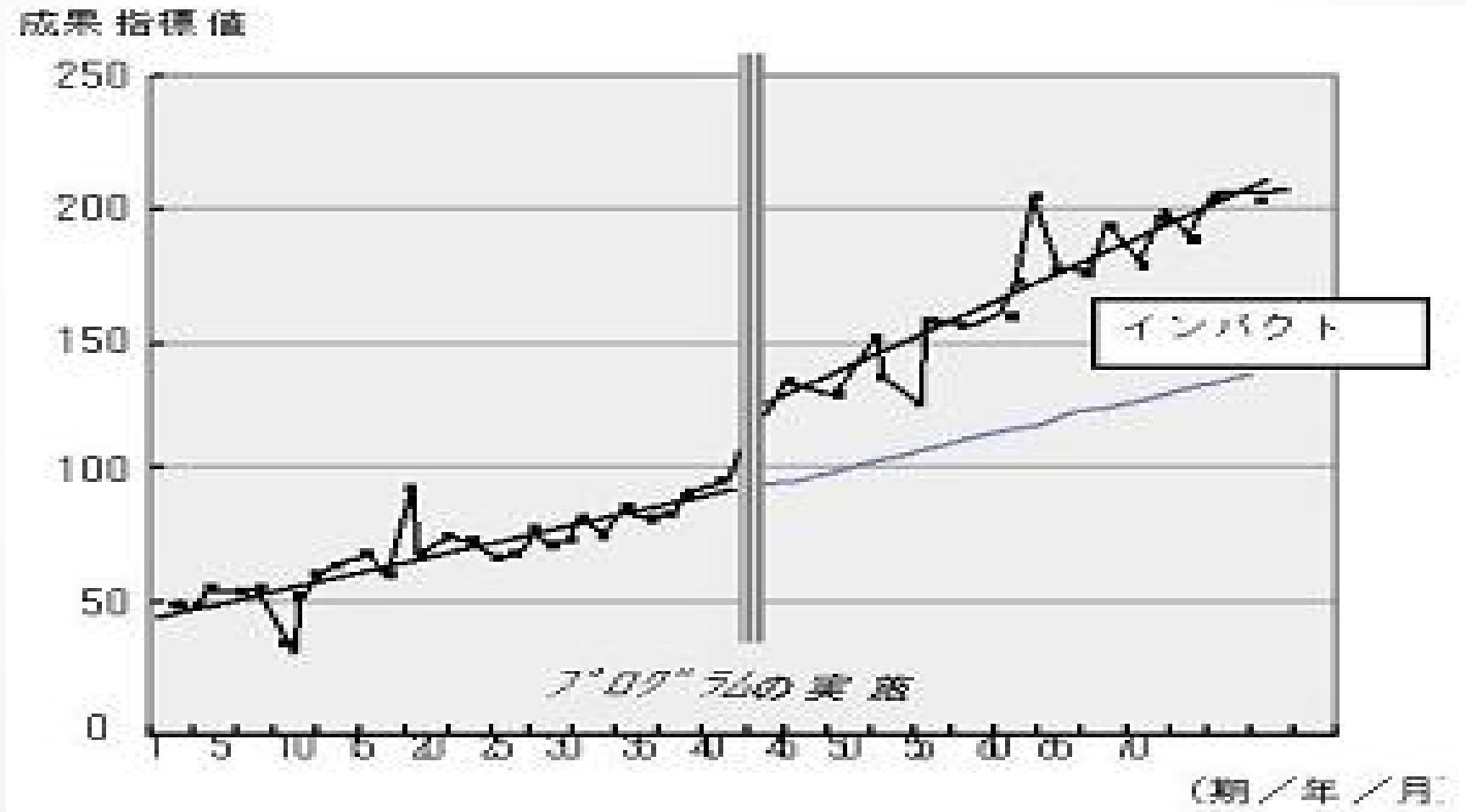
1. 事前・事後比較デザイン

(Before-After Design)



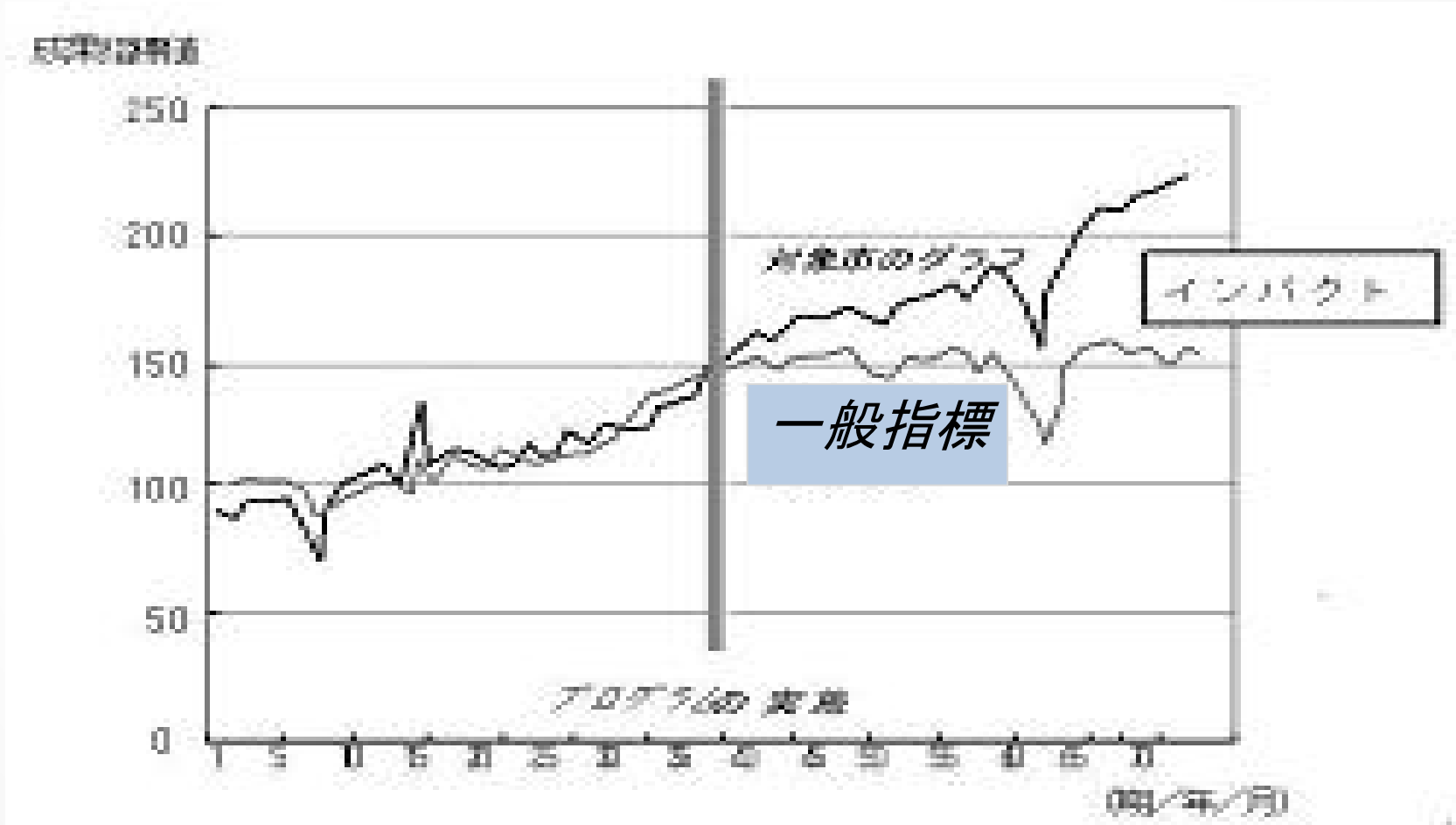
2. 時系列デザイン

(Interrupted Time-Series)



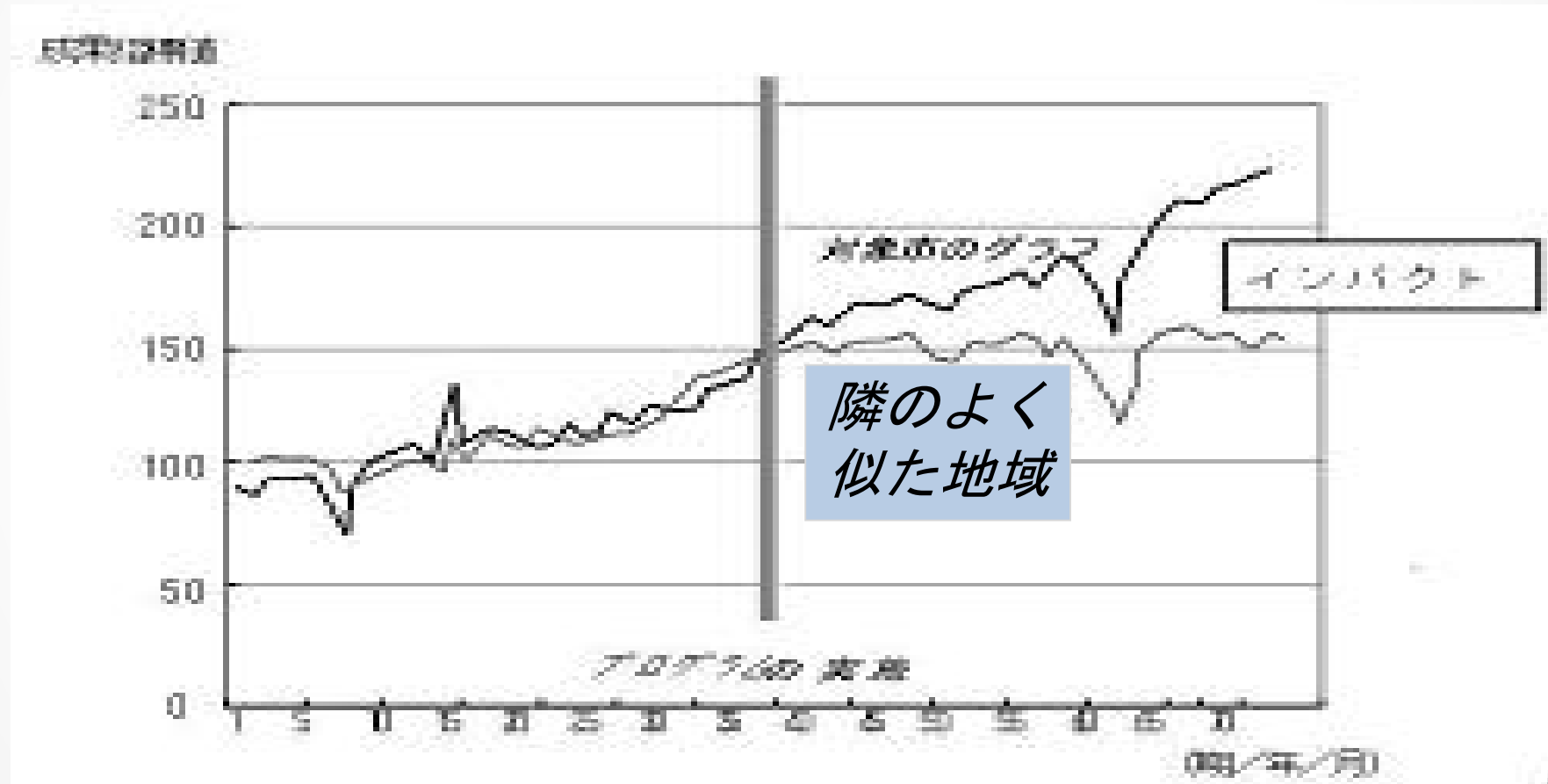
3. 一般指標デザイン

(Generic Control)



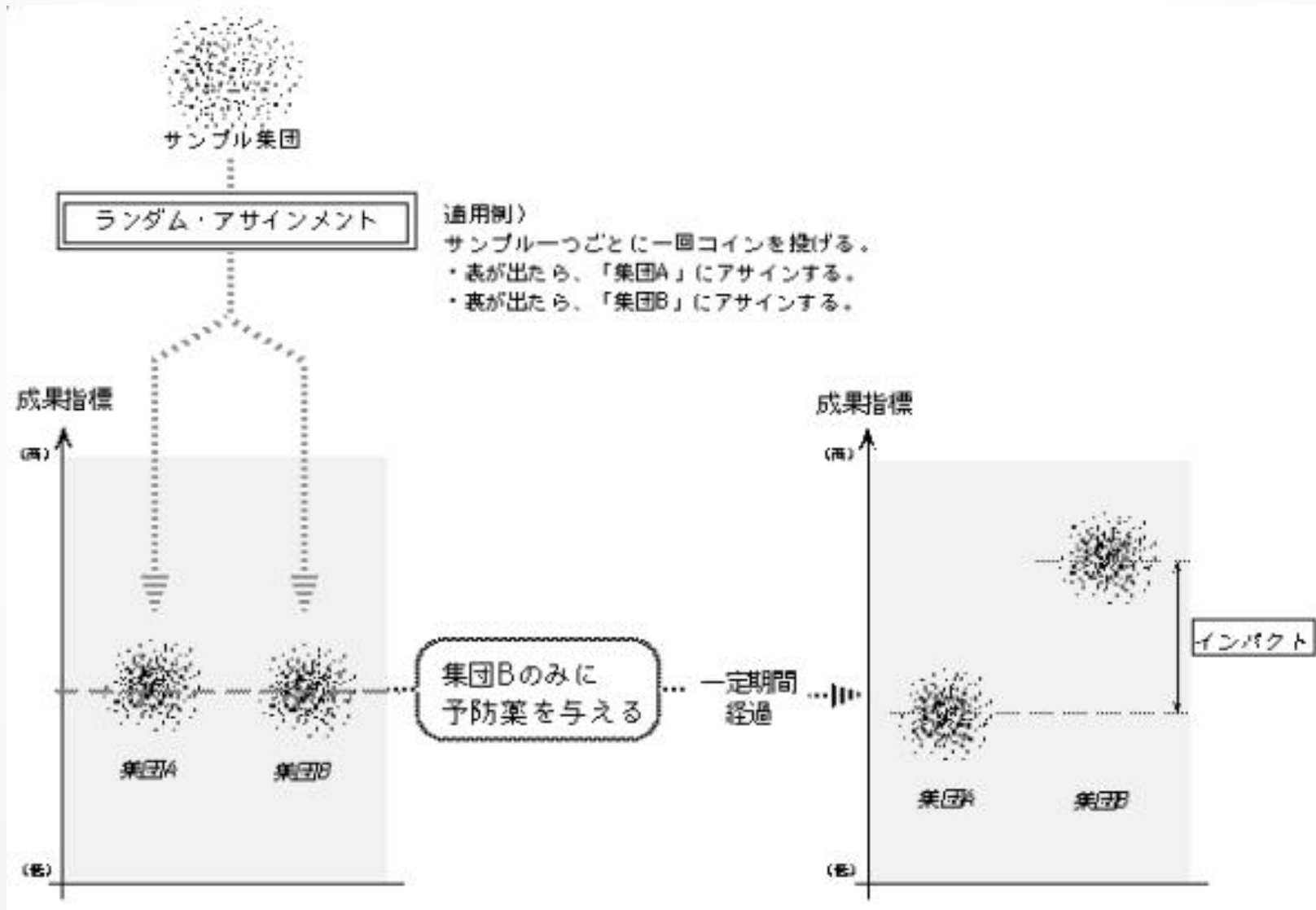
4. マッチング・デザイン

(Matched Control Design)



5. 実験デザイン

(Randomized controlled trial(RCT))



インパクト評価事例集

実験デザイン(RCT)から専門家判断まで 15 の評価事例



Version 5.2 (最終更新日 : 2018 年 04 月 15 日)

評価学博士

佐々木 亮

Ryo SASAKI, Ph.D.

sasaki.ryo@idcj.or.jp

(出所)

https://www.idcj.jp/9evaluation/sub5_files/impact_eval_jirei_28july2011.pdf

業績測定について



Performance Measurement

業績測定 Performance Measurementとは

- 実績測定、業績評価、パフォーマンス・メジャーメントとも
- いくつかの定義
 - ＜業績測定のエッセンスというべき定義＞
 - サービスやプログラムの結果(アウトカム)と効率を定期的に測定すること(Hatry, 1999)
 - ＜業績測定という手法の定義＞
 - インプット・作業(活動)・アウトプット・アウトカム・効率(生産性)などの一連のデータを定期的に収集し報告すること(U.S.GAO 1992)
 - プログラムの実績、特に事前に設定した目標や基準の達成度合いについての体系的で継続的なモニタリング(U.S.GAO 2012)
 - ＜業績測定による評価システム・制度の定義＞
 - 業績(performance)の定期的な測定・評価・対策提言・公表」(上野・上野2007)



Key Performance Indicators (KPIs) の考え方につながっていく

業績測定とプログラム評価の補完関係

- ①手法としての業績測定の基本は、プログラム評価のプロセス評価・アウトカム評価の階層にほぼ相当する。
- ②評価システム・制度としての業績測定の目的は、個々のプログラムにとどまらず、プログラムの集合や機関全体の継続的なパフォーマンス向上や効率化である。
- ③両者は、目的の異なる評価であり、役割分担によって互いに補完すべき(下表)。
- ④業績測定の結果を踏まえてプログラム評価の対象を選ぶ、プログラム評価の結果を踏まえて業績測定の指標を設定するなど、両者の連携も。

Nielsen&Hunter[2013], Hatry[2013]

項目	業績測定	プログラム評価
タイミング	定期的・継続的	定期的でない
評価対象(数)	多数(しばしば全数)を対象に、一律	少数を対象に、1本ごと
情報の深さ	シンプルな業績データ	因果関係を追究
評価体制	内部評価が基本	専門的リソースが必要

業績測定 of 指標に求められるもの

- 妥当性のある指標を設定すること
 - 業績測定において妥当性とはシンプルに「測定すべきものを測定していること」
 - ロジック・モデルに基づいて指標を設定することが推奨されている。
- 目標値は明確に設定すること
 - 根拠となるデータに基づいて、意味が明らかで曖昧さのない数値を
 - 目標達成の難しさが異なれば、達成度合いの評価も異なる。
- 実績値と目標値は的確に比べること
 - 最終的な水準とそこまでの変化のどちらに着目するか、などの的確に選択
- 指標間・プログラム間の比較、集計・集約は丁寧に行うこと
 - 比べられないものは比べないこと(例. フロー指標とストック指標の直接比較は禁物)

おわりに

“評価は社会の改善活動である。”

”Evaluation is social betterment.”
(Michael SCRIVEN)