

政策評価の総論と近年の動向

- I – あらためて「評価」とは
- II – プログラム評価について
- III – 業績測定について

米原 あき
東洋大学・教授
専門社会調査士

I - あらためて「評価」とは

EVALUATION REVISITED

- 評価とは？
- 評価の目的
- 政策評価の目的

「そのもの自体」の“よさ”
内在的価値=本質

評価とは？

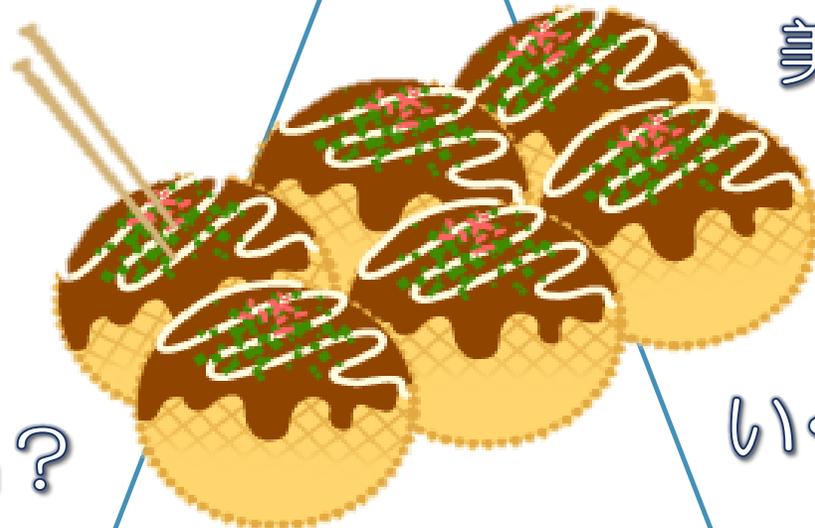
EVALUATION=Ex-VALUe-ActIOn

(「価値value=よさ」を引き出すex活動action)

それが社会（外部）に
もたらす“よさ”
社会的価値=意義

客観的に見た
そのものの“よさ”
外在的価値=値打ち

「そのもの自体」の“よさ”
内在的価値＝本質



美味しい？

有名？

いくら？

それが社会（外部）に
もたらす“よさ”
社会的価値＝意義

客観的に見た
そのものの“よさ”
外在的価値＝値打ち

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

評価の目的：形成的評価と総括的評価

■ 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」

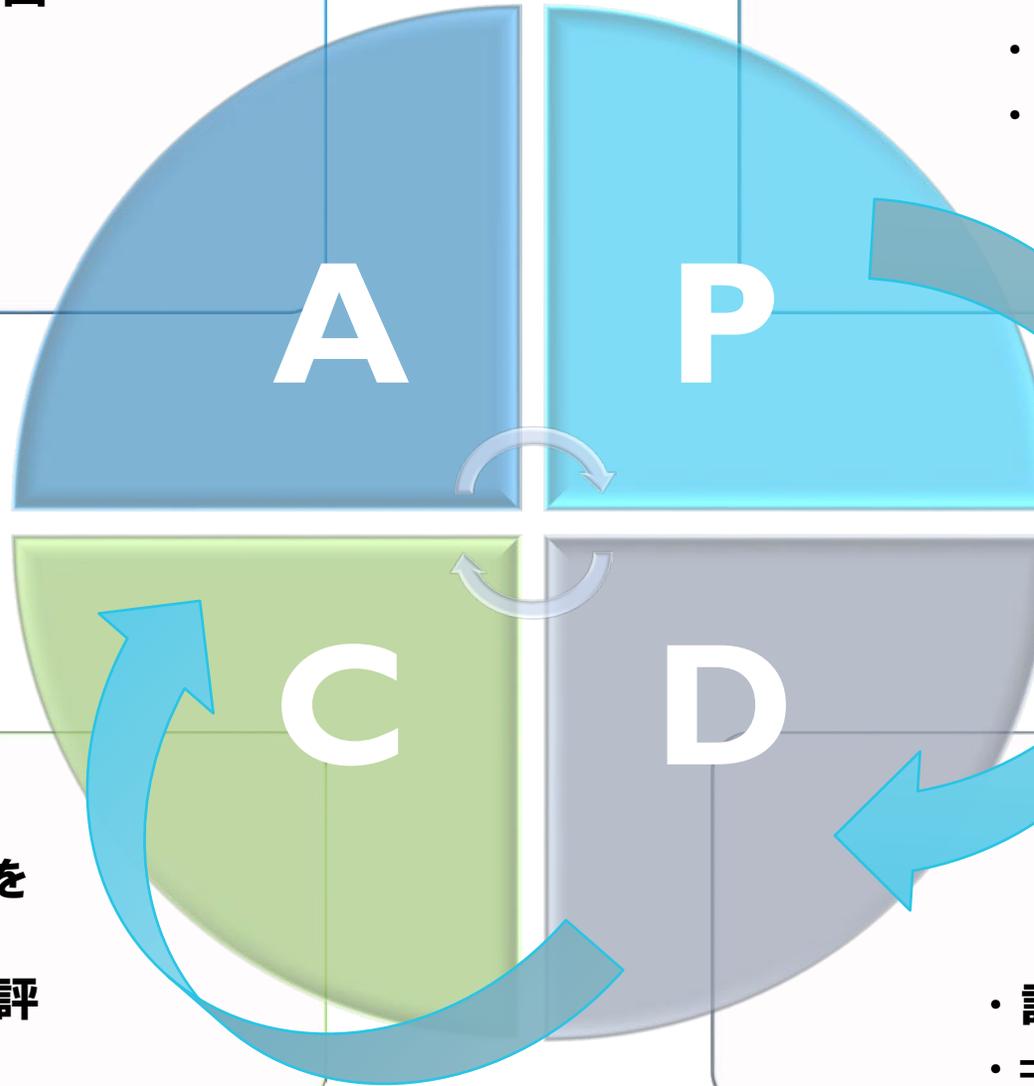
- 事業や活動の成果の把握
- プログラム継続・縮小・拡大の判断
- 一定期間経た段階で行う評価

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

⇒それぞれの特性が生きる
「場面」がある

- ・ 改善への示唆を導出する
- ・ ⇒次のサイクルへ

- ・ ニーズを把握する
- ・ 計画を立てる
- ・ ⇒政策形成の段階



より総括的な段階

より形成的な段階

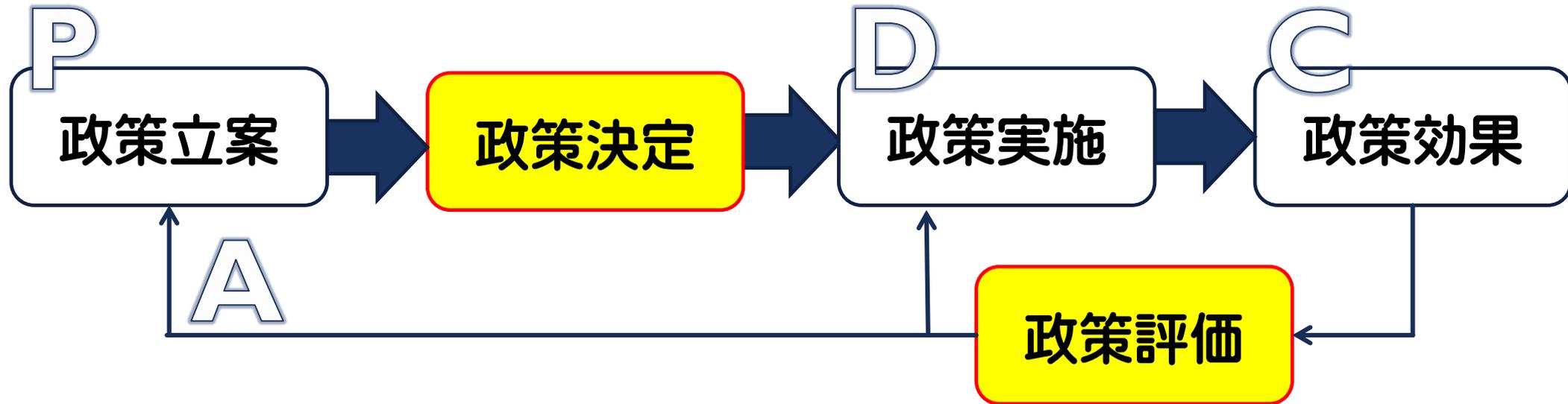
- ・ 実行の結果・成果を評価する
- ・ ⇒政策評価(事後評価)の段階

- ・ 計画を実行する
- ・ ⇒政策実施の段階

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 ≠ 「政策決定」



- 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

政策評価の目的：評価目的の検討

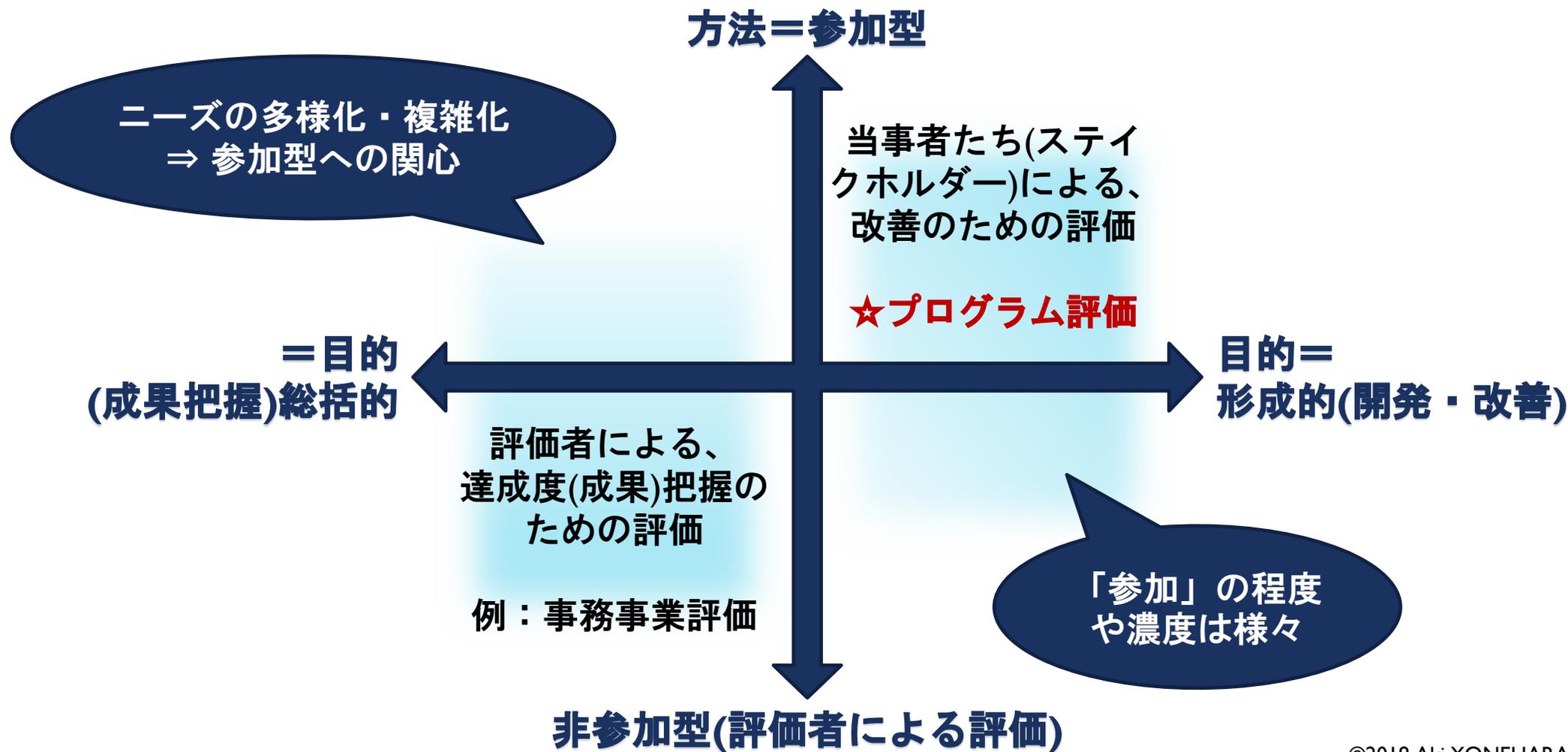
- 事業や評価情報を何のために使うのか（Scriven, 2004）
- 活動の改善を目的としたもの【**形成的評価**】
- 介入のアカウンタビリティの確保を目的としたもの【**総括的評価**】
- 科学的な知識・知見を生み出すことを目的としたもの【**研究目的**】
- など・・・

まず「**評価目的**」を明らかに



目的に合わせて「**評価方法**（誰が、いつ、どのような手法で）」を選択

政策評価の目的：誰が、何のために？



Ⅱ - プログラム評価について

PROGRAM EVALUATION

- プログラムとは
- プログラム理論とは
- プログラム評価とは
- ロジック・モデルとは

何かオカシイ？

STORY_1 健康のための禁煙、ダイエットのための喫煙？

STORY_2 「みどり溢れる活気あるまち」政策評価指標の怪？

目標値KPI	平成X年	平成(X+1)年	平成(X+2)年	平成(X+3)年	平成(X+4)年
公園面積／人 m ²	4 m ²	5 m ²	5 m ²	5 m ²	5 m ²

「健康」 = 体重???

「みどり溢れる」 = 公園面積???

(「活気ある」とは???)

「価値value」を引き出せているかaction? **評価≠測定**

戦略体系図 004

担当部署: 総務部防災課

上位目的: 災害による被害が少なくなる

最終アウトカム

戦略目的: 災害に備え、地域の防災力が強くなる

中間アウトカム

指標: 防災組織の活動状況

→調査→モニタリング

手段01 自主防災組織が活発に活動している

市役所 市民 区 防災士等

0101	組織設立、計画、マニュアル等作成				
0102	出前講座				
0103	ワークショップ、防災訓練				
0104	アンケート実施(活動状況調査、結成への啓発のため)				
0105	自主防災組織の連携				
0106	HPなどによる活動事例の紹介				

指標: 市民の防災知識と意識

→調査→モニタリング

手段02 市民がより多くの防災知識を身につける

市役所 区 公民館 学校 警察

0201	防災学習会、講演会等の開催				
0202	防災マップ等作成				
0203	防災知識の広報啓発				

指標: 防災リーダーの養成・活動状況

→調査→モニタリング

手段03 地域で防災活動ができる人材が増える

市役所 市民 区 消防団 学校、PTA 国 県

0301	防災リーダー養成				
0302	ひょうご防災リーダー講座等の広報啓発				
0303	防災士がいろいろな地域で活動				
0304	こども防災監養成講座の開催				

指標:

有事の際の
死傷者数?



住民の共助
意識!

→調査
→モニタリング

兵庫県豊岡市 『協働型プログラム評価』

参考文献:

源由理子(2014)「地域ガバナンスにおける協働型プログラム評価の試み」『評価クォーターリー』30。

真野毅(2016)「豊岡市における行政評価の取り組み」『評価クォーターリー』37。

米原あき(2017)「セオリー評価における社会調査の活用可能性:『協働型社会調査』の導入事例」『ガバナンス研究』11。



プログラム（PROGRAM）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」**。
- ・ 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。【**未来逆算思考**≡**評価的思考**】
- ・ 「**プロセス理論(過程)**」と「**インパクト理論(成果)**」のふたつの理論が包含される。



Rossi, Lipsey & Freeman, 2004: 140

プログラム評価とは

- 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)

プログラム評価の5階層

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

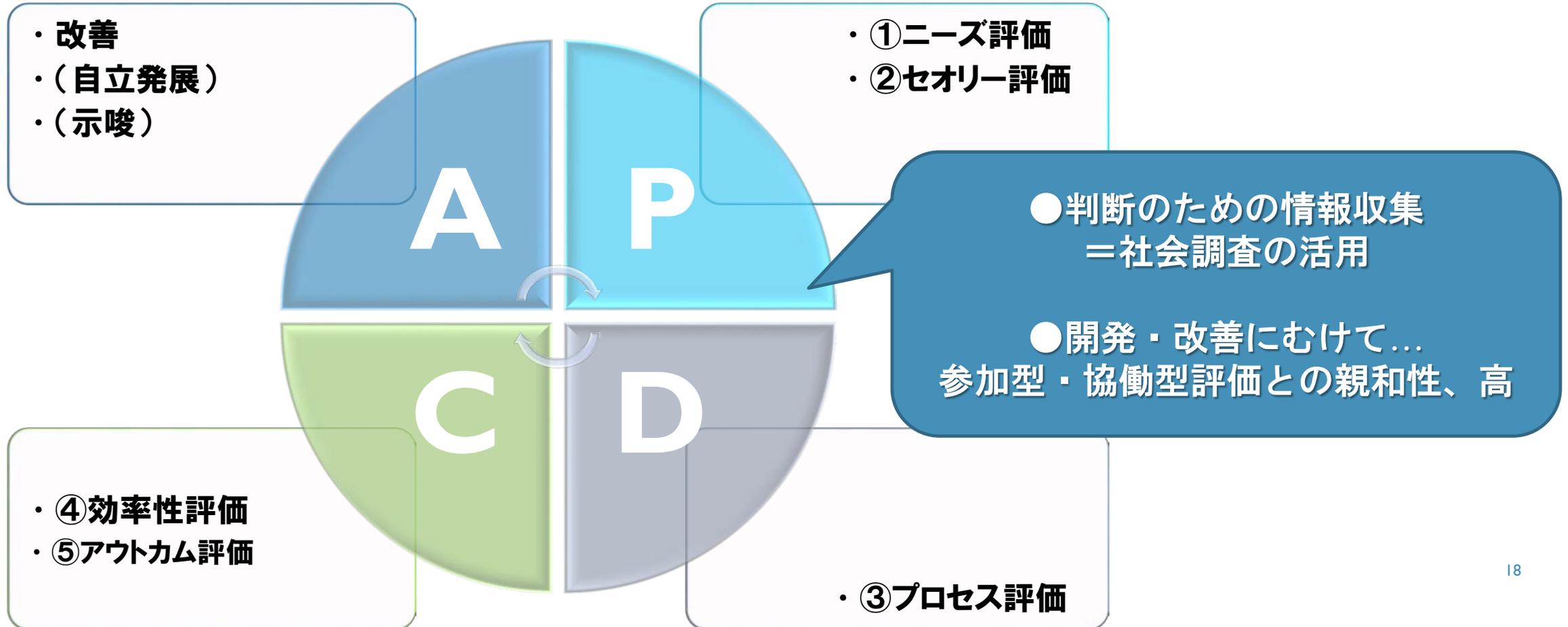
プログラムのデザインとセオリーの評価（セオリー評価）

Assessment of Program Design and Theory

プログラムのニーズ評価

Assessment of Need for the Program

プログラム評価とは



ロジック・モデルとは：定義

■ 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**コミュニケーション・ツール**として活用できる。

ロジック・モデルとは：構成

上位目的(最終アウトカム)：
最終的に実現したい(社会の)「状態」

戦略目的(中間アウトカム)：
上位目的を達成するために実現する
必要のある「状態」

直接目的(アウトプット)：
具体的な取り組みがもたらす直接の結果

活動：
具体的な取り組み

上位目的: 災害による被害が少なくなる

戦略目的: 災害に備え、地域の防災力が強くなる

手段01 自主防災組織が活発に活動している

0101 組織設立、計画、マニュアル等作成

0102 出前講座

0103 ワークショップ、防災訓練

0104 アンケート実施(活動状況調査、結成への啓発のため)

0105 自主防災組織の連携

0106 HPなどによる活動事例の紹介

手段02 市民がより多くの防災知識を身につける

0201 防災学習会、講演会等の開催

0202 防災マップ等作成

0203 防災知識の広報啓発

Ⅲ - 業績測定について

PERFORMANCE MEASUREMENT

- 業績測定とは
- 業績測定とプログラム評価の補完関係
- 業績測定の指標策定の留意点

業績測定とは

- 実績測定、業績評価、パフォーマンス・メジャーメント performance measurementとも呼ぶ

- 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）

⇒ PDCAサイクルの「C」で行う評価活動

⇒ 事後の成果を測定する総括的評価

業績測定とは

■ 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）

業績測定とプログラム評価の補完関係

■ 手法としての業績測定

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。

■ 評価システム・制度としての業績測定

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を目的とする。

プログラム評価の5階層

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（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

業績測定とプログラム評価の補完関係

■ 両者の目的は異なる(下表) ⇒ 役割分担によって互いに補完・連携

∴ 「**違いを理解して適切に使いこなす**」ことが重要！

例) 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 OR

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

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

Nielsen&Hunter[2013], Hatry[2013]

25

業績測定 of 指標策定の留意点

- 妥当性のある指標を設定すること
 - 業績測定における妥当性とは...「測定すべきものを測定していること」
(例：健康＝体重？ みどり溢れる＝公園面積？)
⇒ 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
(例：参加率目標値＝80% ...なぜ？)
 - 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
(例：不登校率＝X%削減 ...学区間の違いは？)

業績測定 of 指標策定の留意点

■ 実績値と目標値の比較の基準を的確にすること

- 「目標値－実測値（目標値への到達度）」OR「介入開始時－終了時（変化率）」

（例：防災研修の受講率）

⇒目標値を達成することが重要なのか、（目標値に届いていなくても）確実に変化が起こっていることが重要なのか、**つまり、上位目的は何か。**

■ 指標間・プログラム間の比較、集計・集約は丁寧にすること

- **比べられないものは比べないこと**

（例：A地区とB地区の自主防災組織・組織率 ...高齡化地域と再開発地域では比較不可能）

（例：教育指標＝入学率、就学率、進級率、修了率、卒業率、進学率 ...似て非なるもの）

27

まとめと要点

- I – あらためて「評価」とは

評価＝事実特定＋価値判断、目的＝形成的評価・総括的評価

- II – プログラム評価について

社会状況を改善するための活動の情報源

PDCAのP段階からはじまる形成的評価

- III – 業績測定について

サービスやプログラムの結果と効率を定期的に測定すること

PDCAのC段階における総括的評価



違いを理解して
適切に使いこなす！

ご清聴、ありがとうございました

【参考資料】プログラム評価登場の背景

- **1950年代** WWII後の米国における社会政策（都市開発、文化教育活動、予防的保健活動など）
➡大規模プロジェクトに対する説明責任の発生 ➡評価への関心
- **1960年代** 米国の大規模社会政策（e.g. ケネディ、ジョンソン政権「貧困との戦い」、「偉大な社会」）
➡“科学的な”評価研究の発展
- **1970年代** 行政分野における政策分析との結びつき
➡“科学的な”評価研究を通じた「費用便益分析」や「財政的説明責任」
- **1980年代** 税負担と公共サービスの質に対する市民の不信
+ レーガン政権の赤字縮減政策＝社会プログラムの予算削減
➡**プログラム評価（対人社会サービスの評価）の発展と普及**
（科学的評価への懐疑と代替的手法の提案）
- **1990年代** 様々な代替案の登場・・・参加型評価、エンパワーメント評価、実績測定
- **2000年代** 新たな潮流・・・科学的評価の復活Evidence-Based Practice/Policy
➡**評価の世界的普及と目的の多様化**