

政策評価に関する統一研修（地方研修）那覇会場講演概要

平成 27 年 7 月 22 日開催

講義名：客観的かつ具体的な政策効果の把握及びそのための適切な目標・測定指標の設定について（演習）

講師：鳥取大学地域学部教授 小野 達也

講義時間：13 時 10 分～14 時 40 分

テーマ

→ 今日の演習テーマは、「客観的かつ具体的な政策効果の把握及びそのための適切な目標・測定指標の設定について」であるが、サブタイトルとして私の方で、「～成果指標と目標値の効果的な活用に向けて～」と付けている。

演習型研修ということであるので、話を進めて行く中でところどころで時間を取って、具体的な問題として皆さんに考えてもらうようにしたい。その後で、それについて解説したい。

→ サブタイトルは、指標と目標値によって政策（政策・施策・事務事業）の効果を客観的に把握すること、それに関することである。今日は政策評価の統一研修ということであるが、地方自治体の方もたくさんおられる、国の機関、自治体、独法やその他いろいろな公共機関に共通的な話であるので、自分に関係する部分に引き付けて考えてもらいたい。

指標とは、

→ 原則として、数値で測定する指標。

・呼び名は政府の政策評価制度では測定指標、自治体では呼び方は様々で評価指標、業績指標など。数値による把握が不可能であれば、文言で記述する指標を用いることになるが、これは後から触れたい。

把握の対象となる効果とは、

→ 成果（アウトカム）は何か、政策の成果は、政策をどれだけ行ったか、行政サービスをどれだけ提供したかということではなく、その結果として何ができたか、できなかったかも含めて成果を把握したいということがここで言う効果である。成果（アウトカム）の把握が難しい場合はどうしたらいいのか、現実的には重要であるので話の中で触れていきたい。

プラン

→ これからの説明は、1. 数値の指標を使う意味を述べて、2. 次に、各論として、4 つ位に分けての説明。そのところどころで具体的なケースの話をし、皆さんに演習という形で演習問題を具体的に行ってもらい、解説し、3. その後で、Q&Aの時間を取りたい。

1. 数値の指標を使う理由

○ 評価において、数値を使うメリットは3つある。

①客観的に表現できる。文言・言葉でも表現できるが数字は客観的表現になりやすい。数字はウソをいわないはず。

→ 1,000 より大きければ大きいと言うなどのルールが決まっていれば明確である。大きいか大きくないか判断ができる。表現の使い方ではっきりする。

②比較や加工ができる。

→ 数字で表すことにより昨年より大きくなったか、高くなったか、良くなったか、

隣の自治体と比べてどうか、日本国内都道府県別に見てどうか、平均的に過去5年間でこれくらい平均何%増えたかなど、比較や加工ができる。これは数字ならではであり、比較することによって数字は多くを語る。

③わかりやすい。

→ 数字は分かりやすい説明となり、意味がはっきりする。明確に意味が伝わる。

何か大きいと言ったときに数字で示せば大きいということが分かりやすい。文言で説明することはわからないというのでなく数字で示せば意味がより伝わりやすくなる。

※1 実は評価に限らず世の中での数字を使うメリットはこの3つになるのではないか。

※2 どうしても数値化できない文言による指標にも、数字が持っているメリットのこの3条件に近いものが求められる。文言であっても客観的にするべきであるし、ある程度比較が出来なければ使いようがない。分かりやすさも求められる。

○ただし、評価における客観性には2種類ある。

個々の主観から独立であること。(普遍性)

→ 数字で測定し、把握するということは何を意味しているかというところと誰が把握しても同じ結果になるはずであるというところにある。そういう意味で個々の主観から独立している。

特定の立場にとらわれないこと。(不偏性)

→ 評価の場合、特に内部評価と外部評価は典型的。内部評価は意図的でなくてもどうしても甘くなる。それを外部で補う。そういうときに客観性ということになるが、実は、数字を作る場合もこの話は無縁ではない。数字をねつ造することはなくても都合のいい数字は使い、都合の悪い数字は出さない。意図的な場合もあるが、図らずも、上がっている数字が出ていて下がっている数字は出ないこともある。

→ 評価における客観性はこの2つがあり、数字もそれぞれ関わりがある。

○数値指標のメリットと裏腹のデメリット

→ 客観的に表現できるが数字になっていることで客観的に見えてしまう。評価指標なども典型的にそのリスクがある。

→ 根拠が定かではない指標だが評価指標として書けると、それに基づいて何とかということになると、数字にすると客観的に見えてしまう危険性がある。比較や加工ができるが、おかしな計算をすれば元も子もない。そもそも数値自体がきちんとしていないと比較する意味も確からしさもない。

→ わかりやすいがゆえに数字が独り歩きする危険性もある。評価指標もそうで、メリットがあると同時に使い方には気をつけなければいけない。

○評価(業績測定型評価 Performance Measurement)

→ 評価指標は、評価においても客観的にということであるので、数字の指標を使うということはいろんな局面で求められている。国の政策評価制度でも自治体の評価制度でも、先行している自治体では、国によりも前に古くから行なわれているが、どこにおいても一定程度は数字の指標を重視するというのが多い。今日は、業績測定型の評価ということで、指標や目標値で目標管理をする、あるいはPDCAサイクルを回していくというのは、評価の分野ではパフォーマンス・メジャーメント(業績測定)と言われており、広く浅く全施策、全事業について評価対象にし、指標を設定し、目標値を設定し、進行管理をするものである。

- 参考資料の事前分析表は、午前中の説明にもあったが政府の評価で中心的な役割を果たす各省庁の主要施策全てについて対象となっている目標管理型の政策評価において記載するものである。上の方に達成すべき目標があり、その横に目標を設定した考え方・根拠を書く欄があり、その下に測定指標の欄があり、その右側には測定した根拠どういう理由で選定したか。目標値はどういう根拠で設定したかを書くことになっている。
- 次の参考資料は、評価書のフォーマットであり、評価指標を中心に測定指標の欄があり、それに基づいて分析・評価するものである。次にある資料は実際の例で、農林水産省の実際に記入された例である。農林水産省は数値をきちんと積極的に使っており、データが豊富にある政策領域が多いので挙げさせてもらった。次の資料は日本で最初に体系的に行政評価に取りくんだ三重県のインターネット上で公開されている一番古いもので平成 10 年度の例であり、事務事業評価と言われているものである。事務事業の目標がきちんときめ細かく記述されていて、指標の設定、定義・説明もあり、指標の目標値、実績値もある。数字以外の分析もあるスタイルで自治体の典型的なスタイルとして広まった。次の資料は岩手県の事例であり、ほとんどの自治体で作っている総合計画の進行状況を数字を駆使して、きめ細かくチェックする。指標 1 つについて 1 枚のフォーマットにきちんと説明している。
- 自治体でも国でも評価指標と目標値を設定して、PDCA サイクルを回していくというのはあらゆるところでやっている。
- 行政事業レビューもお金の使い道、ムダをチェックする役割もあるが、事務事業単位で目標管理をする役目も持っている。地方創生は沖縄県内でも市町村・県庁も含めて総合戦略を作られていると思うが K P I を作れ、成果（アウトカム）を測る指標を設定してきちんと達成度評価をして下さいということになっているので、行政評価制度とは別に評価指標を作らなければいけない場合も多いだろう。
- 今日は、共通する話として、指標と目標値の話をしていきたい。

2. 成果を測る指標と目標値を設定する。

具体的な各論に入っていきたい。4 つに分けてお話をしたい。

- (1) 達成すべき目標、目指すべき成果を明確にすることが大前提にしてある。
 - (2) 2 番目に妥当な指標を設定することが重要なキーワードとしてある。
 - (3) 指標を設定した後で、目標値を設定する。目標値も妥当でなければいけない。この妥当というのは限定した意味で使っており、また後で説明したい。
 - (4) 最後に測定（記録、計算）の信頼性である。軽く触れるだけにしたい。
- (1) 達成すべき目標、目指すべき成果が明確になっていること。
- 達成すべき目標、目指すべき成果を明確にしておかないと、そもそもアウトカム指標までいかない。これは当然ではあるが、国の政策評価制度においても反省点として、午前中の話にもあったが、その辺りが、導入後 10 年以上も経っているが必ずしもうまくいっているとは言えない。自治体の行政評価でも似た側面がないわけではない。
 - ロジックモデル。政策の成果が発現するまでの過程をロジック（論理）という。ロジックは、インプットは、行政の投入、予算や職員の労働時間も含めて行政資源の投入であるが、そこからアウトプットという形で行政サービスやハコものなどいろんな形で、国民・市民に提供される。その結果どういう成果があったのかを段階に分けて考えていきたい。

→ 例として、小中学校にパソコンを導入する（PC 導入事業費）→学校に PC が導入されて使われるのがアウトプット→PC が活用されるのが成果。導入されたが活用されないと、成果がないということになり、アウトプットはあるがアウトカムがないということになる。

○政策のロジック（詳細版）

→ 今のはインプット、アウトプット、アウトカムの3分類であったが、もう少し細かく、きっちりと評価をしようということになると、インプットとアウトプットの間にはプロセスがある。行政機関・公共部門の中でアウトプットを出すまでの間に仕事がある。パソコンを導入しようとするれば、計画やもの手配、設置する作業などがあって学校に PC が導入される。アウトカムは1番重要と考えて見ると、きめ細かく考えると3段階に分けられる。どこからどこまでが直接アウトカムで、中間的アウトカムはどこからでと決まっているわけではない、順序関係である。PC が導入されて授業で使われる状態になる。しかし、これで終わりではない。最終的な目標はもっと先にある。教育効果が向上する。何らかの効率化をすることにある。児童・生徒のパソコンを使う能力が向上し、より効果的にパソコンを使った授業ができる段階があり、最終的には教育の効果が上がるということを実現したいということにある。

○「少子化対策の施策」のロジック・モデル

この中のどこをどう測定するのかということになると指標の設定の話になる。

→ このロジックモデルは私が関わったものに加工したものになるが、少子化対策として、出生率の維持回復をしたいことがあり、そのための事業、施策がある。最前線では補助金などで基礎自治体の中でいろいろな制度でいろいろなことがなされる。これは国の例となるが、地方公共団体でも同じであるかもしれない。例えば、子育て支援の話になるが、育児休業制度→制度の充実→育児休暇が取得されているか→育児と仕事が両立しているかとなるが、そのブロックの下にはそれぞれ、指導件数→制度の普及率→取得率→両方できると感じる割合が指標として入っている。

→ ロジックモデルの流れとしては、アウトプットが入り、アウトカムへとつながる。アウトプットとして「制度が充実」し、直接アウトカムとして「育児休暇」の取得がどの程度されているのかとなり、中間アウトカムとして「育児と仕事の両立」がどの程度なされているのかとなり、最終的アウトカムとして「子ども数の回復」につながる。保育所の増設、規模拡大についても同じ流れである。保育所の定員が増加につながり、アウトカムとして待機児童数が減り、仕事が続けられることになり、最後は出生率の回復となる。ただ、ロジックモデルでは、子育てを支援することとは別の要因もあり、日本の出生率の低下は3分の2ぐらいは結婚の問題であり、結婚時期が遅くなったり、結婚しない人が増えてきたことに理由があると言われているので、少子化対策を一生懸命やっても限定的な効果しか出生率にはないということも明らかになっている。そういうことも含めてこのような論理的関係を整理することがまず必要である。世界的にもそう言われている。全部の事業や施策についてこのような絵を描くのかと言われると、確かに大変であるがそういうことを行っているところもある。日本でもそうしたことをしなくてははいけないという機運は徐々に高まってきている。綺麗な絵を描かないまでもこうした考え方は基本的に指標の設定や目標値の導入と言ったところで必要であるということになる。政府の政策評価制度でも午前中触れていたように、

ロジックという考え方をして、達成すべき目標を考えたり、指標を設定して下さいということに成りつつある。

□達成すべき目標、目指すべき成果の記述例

→ 達成すべき目標、目指すべき成果の記述例ということで、現実的に使われている例と加工した例を挙げている。○印はよい例、問題のない例であり、×は問題のある例である。

○ 意欲ある多様な農業者による農業経営の育成・確保

→ かなり具体的に書いてあり、どういう指標で評価するのかということはいろんな指標をある種思いつくだろうというものである。

○ 事務事業評価システムの導入により職員の意識改革や能力開発が進み事務事業の改革が進められる。

→ これは直接測る指標は、実際には、工夫しなければ考えられないかもしれない。三重県の例であるが、具体的に工夫して測定する指標を使って書かれている。

× (成果を評価することを謳っているが、あるいは成果指標を設定しているのに) 目指すべき成果の記述がない(記述欄がない)。

→ 成果を評価すると謳っているが、目指すべき成果の記述がない。あるいは、成果指標はあるが、何の測る成果指標か分からない。指標の欄はあるが事業のある施策が何を指すのかどこにも書いていない。明確になっていない。

× ……のために○○対策を着実に推進する。

→ これはしばしばある表現であるが、国の政策評価制度の事前分析表を見ても少なくない。自治体の評価表を見てもこのような表現は少なくない。これは、成果のことを言っていない。アウトプットにもなっていない。プロセスをきちんとやると言っているだけ。成果とはほど遠い。

× ……について、課題の検討を行うとともに施策を実施する。

→ これもアウトプットにもなっていない。成果指標は難しい事情があるとしても、評価をする仕組みからすればダメである。

→ 具体的に書かないと、指標を設定するところまで本来いかなければならないはずである。事業でも政策でも何を指すのか、何が実現したいのかがそもそも明らかでなく、ロジックモデルの右側の方がぼやっとしているままでは、指標以前の段階で指標に進めない。本当はあるのだが説明していないということならば評価情報に入れるべきということになる。

(2) 指標が妥当であること

→ これは指標を設定するときに、まず求められる。この妥当性というのは社会科学の用語としてはvalidityという定義があり、測定すべきものを測定していること。

→ 欧米の方での譬え話として、私たちが真っ暗の夜道を歩いていて、コンタクトレンズやピアスなどを落としてしまい、真下は真っ暗なので探すことができないときに、少し離れたところには明かりがあって地面が照らされているが、そこで探すということには意味がないことになる。この譬え話の意味は、あなたが使っている指標というのは、そういうものになっていないか。手近なデータがあるのか思いついた指標を使っているというのだが、実は自分の真下で探していないのではないかとということになる。本当に測定すべきものを測定していないのではないかとということがこの妥当性ということである。

○指標の妥当性を決める重要なポイント

→ まず重要なのは、ロジック上のどのブロック（ロジックのハコ）を測定・把握の対象とするのか、アウトカム（成果）ということであれば、ある事業なり、政策を行ってこういうことを世の中に提供すればこういうことが起きるはずである。パソコンの例でいうと、パソコンを配置し、この段階での評価指標ということではパソコン1台当りの生徒数が何人実現できたということの意味のある指標となっているが、これはアウトプット指標である。アウトカムの指標では、パソコンが授業で使われるようになる。これは実際に使われている指標であるが、これは絶対に悪いというのではないが、本当に目指すところではない。中間段階では教師の習熟度があり、うまくパソコンを使う段階がある、何らかのアンケートでそれを把握できるかも知れない。最終的には教育効果が向上したり、効率化することにあるが、これを測定することは簡単なことではない。何らかのテストやそれに類するものをやらないと、事前事後と比較をしないと本当のところどれだけ効果があったかは分からない。事後的にアンケートを行って評価することは可能かもしれない。一概にどれが正解かは言えない問題である。どれぐらい評価に手間を掛けられるのか、あるいは時間を掛けられるのか。どれだけ力を注いでやるのかということがある。評価疲れということもあり、あまり評価のために時間を掛けると本業を圧迫することもあり、本末転倒ということになる。できる範囲で行うことが評価の基本原則だが、何が一概に正解とは言えない。あまり余裕が無くて、学校に協力してもらい調査をするぐらいかもしれないし、もう少し時間を掛けられるならば、もう少し踏み込んだ調査を行い、どの程度効果があったのかを測定すべきであるのは当然である。

<演習>

Q1 ここに3つ政策がある。アウトカム（成果）に着目して評価を行う場合に、どの場面・段階（ロジック上のブロック）について、指標と目標値を設定すべきか。どのマスを皆さんであれば設定すべきと考えますか。次の3つどれでも皆さんが関心のあるものについて、1つ選んでロジックモデルを考えて、どこを評価対象とすべきなのかを考えて書いてもらいたい。

- (1)子育て支援策（育児休暇制度の利用促進、保育施設の充実など）
- (2)地域の人口Uターン誘導策（情報発信、相談機能の充実、空き家活用など）
- (3)自治体のプレミアム（付き）商品券発行

※ 最後に解答をまとめたものをお配りするので、スライドを見てもらいたい。
解答はこれしかないというものでなく一般的な解答ということで、教科書的な解説ということで聞いておいてもらいたい。最終成果に一番近いものが望ましいが、あまり最終成果に近づくと他の外部要因が絡んでくる。評価対象の固有の成果と言えるものなのかという観点もある。これは必ずしも両立しないのでどこかでバランスを取ることが必要である。

A1 Q1(1)

- ・「育児休暇制度」については「育児休暇の積極的取得（指標：取得率）」が有力。一番代表的な指標という意味では、育児休暇が積極的に取得されればいい。
- ・「保育施設」については「入所待機児童の減少（指標：待機児童数）」が有力。
- ・「(働く親のための)子育て支援策」については「育児と仕事の両立（指標：両方できると感じる割合）」が有力。これは何かアンケートをしなくてはいいな

い。

Q1(2) Uターン誘導のために情報発信する、相談体制を整備する、空き家バンク発足させるがあり、それぞれ対象となる方に、情報が到達する、相談が増加する、空き家バンクが登録されて問い合わせが増えてくる。それが実際にUターンに結びつく、Uターンが増加して最終的に人口減少を抑制したい。

ポイントとしては、最終成果に近く、かつ政策固有の効果として把握可能なことが挙げられる。

Q1(3) プレミアム付き商品券が販売されたり、抽選だったりするが、商品券が使われる、地域消費喚起・生活支援という側面もあるかも知れないが、短期的経済効果がどの程度上がったのかが最終成果に近く、かつ政策固有の効果としてタイムリーに把握が可能なアウトカムとなる。最終的な地方創生ということになると他のいろいろな要因が入って来るので、プレミアム商品券が使われている時期に地元の商店街でどれだけ売り上げが増えたのかということが、何らかの形でそういう観点からの評価が行なわれないとやった意味がつかめなくなる。

○指標の妥当性を左右する条件

指標の妥当性を左右する条件として8項目に整理している。

①～④までが指標を設定する本質的な問題、⑤以下が実務上のテクニカルな問題

①目標との適合性：ロジック上の位置が明確か

→ きちんと成果を測っているのかとか、ロジックがつながる意味のある目指すべき成果として明確になっているものを指標化しているか。目標との適合性と書いている。これが一番重要と言えれば重要である。

②包括性：政策の重要な側面をカバーしているか

→ 政策のある一部だけでなく、施策なら施策の全体をカバーする代表的な指標になっているのか。対象が狭すぎないかということである。

③外部要因の影響：他の政策、他機関の政策、行政の外の要因などの影響がどれくらいか

→ こちらは逆に広すぎないかということで、ロジックモデルの最終的なものにするのはいいいが、子育て支援策をいろいろして合計特殊出生率の向上という、結婚のことも大きく関わってくるし、あるいは自治体であれば県の施策であれば市の施策が関わってくるし、国の政策が関わってくることもある。他のものも関わってくるのであまり野心的に最終的な成果を目指せばいいというものでもない。外部要因が入ってしまうと評価対象の当該政策の評価と言えなくなってくる。

④感度、感応度：効果（結果）の大小に指標値が反応するか

→ 具体的にどのブロックをどの段階の成果指標とするとしても指標自体がそのブロックを測るものとなっているか。

⑤計測可能性：実際に安定的に測定できるか

→ 実務的なものであるが、概念としてはあるが実際に安定的に測定できるか。

⑥適時性：タイムリーに測定できるか

→ 時間が掛かる、時間が経たないとわからないとか、周期的に行なわれている国の統計調査の結果は使えるか。

⑦費用：データ収集コストが過大でなく、予算内

→ アンケートをする。お金が掛かる。評価にどの位お金を掛けるのか。政策の

効果をどこまで評価すべきなのか次第になってくる。午前中の質問にもあったが、教科書的にこれですという解答はない。予算内の範囲内でできる限り行うべきであるということにはなる。設定すべき指標を設定すべきである。

⑧悪影響の可能性：局所への集中、データ操作などの可能性

→ その指標に関わるところだけに集中してしまうとか、どうしても評価指標をよくしたいということになると偏ってしまう。もっと悪いのはデータの操作である。日本でもメディアでも報道されたこれに近いこともある。

□成果指標の設定例

○研修参加者の知識・技術の修得度合（事後アンケート）

→ 職員研修をする、どれぐらい知識・技術の修得ができたかを事後のアンケートで取る。本来標準的なよくあるパターンである。

×研修参加者の満足度（研修終了時のアンケート）

→ 一方満足度を聴く、大学でも授業の満足度ということで、学生の授業評価ということで、最近はそのを教員の評価の重要部分と言われて、危機感を抱いているが、本来あまりよくない。授業評価は別として、研修参加者の満足度は全く意味がないかは別として重要な指標ではない。満足したというよりも身に付いたかどうかである。

○刑法犯の認知件数→通常使われており、意味がある。

×凶悪犯の検挙率→△かもしれないが、凶悪犯を検挙する。アウトカム指標というよりもアウトプット指標である。凶悪犯にターゲットを当てても凶悪犯罪が起きないようにするということを目指すべきである。凶悪犯に限定せず、安全な地域とか安全な社会ということであれば、凶悪犯だけを代表的な指標とするのはどうかということになる。日本ではないが、凶悪犯が野放しになっていて、検挙率が低くて、何とかして検挙率を伸ばすということであれば立派なアウトカムになるかもしれない。

○（自治体域内の）製造業の自立度（域内需要の生産誘発額、産出先・投入元の域内外比率）

→ 工夫はいろいろと必要となるが域内でどれだけ自立しているのかということ測定可能である。

×（自治体域内の）製造業出荷額

→ なぜ×なのか。これは製造業は典型的にそうであるが、日本の自治体の場合域内で完結していることはまずない。原材料や中間財をどこから持ってくる。あるいは出来上がった製品を中間財としてどこかに提供する、あるいは消費者に売る、いずれにしても域外にも相当出回ることになる。労働力も域外から調達しているだろうし、いろんな金、物、サービスの流れも域外とのやり取りが相当程度あって、製造業は成り立っている。日本国全体でもそういう側面があるが、自治体ということになれば、自治体の境に関税が掛かるわけでもないので基本的に自治体の中の製造業を、市や県が支援してどうなったのかという評価には、製造業の出荷額はあまり適当ではない。外部の影響があまりに大きすぎる。域内ということを確認にしないと評価指標としては苦しい。

○/×合計特殊出生率

→ これは自治体であれば、男女の出会いを増やす取り組みを増やすとか、結婚なども含めていろいろやっており、最終的に出生率がどうなのかということであればよいが、子育て支援の策をいろいろして、合計特殊出生率がどうなのというにはその間には距離がある。結婚の問題は大きいので、この指標は使っていけないことはない

が、これで代表しますということでは適当ではない。

○/×待機児童数

→ これも同じで、保育所の定員を増やして、まさにそれを測る指標となるが、先ほどのロジックモデルの例でその先がどうかということに重点を置くと、他のことを含めて子育てが安心してどの程度できるようになったのかということであれば、待機児童数だけでは狭いし、手前すぎることになる。

○/× A 事業の経費削減額

→ ある事業の経費削減額が指標になっていたりする。行革のプログラムのようとにかくできる限りスリム化する、削減自体を目標に取り組む、それ自体が事業であるということであればそのものずばりとなる。ときどき事業の成果指標として経費を削減・節減しましたと言ったときに、それはアウトカムではない。一方成果は変えずにコストが削減できたということであれば、効率がよくなったということなので評価指標にはできるが、削減額自体は事業のアウトカムではない。

○勧告に基づく改善率 ×勧告件数

→ 勧告をいくらしめたというのはアウトプット指標となる。勧告をした結果どの程度改善したのか。勧告した対象の 95%は指摘に基づいて改善したということはアウトカム指標となる。

△施策の推進（具体的記述あり一項目・件数・時間など）、×施策の推進（具体的記述なし）

→ 項目・件数・時間など測定指標として、具体的記述があればまだしも△であるが、施策の推進という記述だけなのはダメである。

<演習>

Q2 次の指標は、何れも成果（アウトカム）指標としてしばしば用いられているものである。これらの指標の妥当性は十分でしょうか？

（○：妥当である、×：妥当でない、△：○と△との中間。判定だけでなく理由も考える。）

(1)男女共同参画政策の成果指標としての「審議会委員の女性比率」

(2)雇用対策の成果指標としての「有効求人倍率」

(3)不法投棄対策の成果指標としての「摘発件数」

A2 Q2(1)

・結論的には×である。男女共同参画の成果指標としては幾つかあり、そのうちのひとつとなっており、そのうちの一部にしかすぎない、外部要因も考えられ、比率のルール化（審議会の委員の3割は女性にしなければいけないとかのルールがあり、そのため3割確保するために努力する、など）があればそれはアウトプット指標である。）

Q2(2)

・雇用対策をいろいろして有効求人倍率を見る。これは最終成果に近いと言える。いろんな取組をして有効求人倍率が高くなるのはいいが、外部要因は世の中の景気の動向、自治体から見れば国の政策、民間機関の景気、あるいは外国のリーマンショックなどのようなことも関係してくるかもしれない。外部要因も大きい。その他にこのブロックを測定する労働市場の需給関係が改善することにしても、ハローワークを考えると少し狭くなる。正確な数字は覚えていないが、職を得る人はハローワークが大部分を占めているわけではない。3割ぐらい

かもしれない。就職情報誌とか他のチャンネルがいろいろある中でハローワークの指標はある程度限定的になる。△ぐらいになる。カチツとした統計指標となるので使い安いと言えば使い安い。

Q2(3)

・不法投棄の対策事業の中で、いろいろ取締をしたり、警戒をしたりする中で成果指標の摘発件数ということであるが、これは○の場合も×の場合もある。両面性がある。

・(ケース1) 対策に着手したばかりである、摘発自体が困難な取り組みである、などの事情があれば、成果として意味があり、○。

・(ケース2) 一方、既に一定期間、継続的に対策に取り組んでいるのであれば、摘発件数はアウトプット指標であり、×となる。

(3) 目標値が妥当であること

○目標値の妥当性を左右する条件には3つある。

①いつまでに、どれだけの水準を目指すのが明確であること

②目標値の性格が明確であること

例えば、西尾勝先生がアメリカ行政学を学んで書かれたものであるが、

i)理想状態を示す期待値

ii)一応の水準として満足できる充足値

iii)絶対達成すべき限界値

→ この3つがあったとして、例えば目標値、目標には未達だったけれどぎりぎり近い線までいきましたという実績値になったとして、目標値が上のどれかによって意味が全然違ってくる。i)の目標値であればこれはよくできたことになる。一方最後のiii)であれば、近いと言っても実現できていないので何が問題であったかきっちりと考えなければならなくなる。これが明確でない場合が結構あり、性格がはっきりしていないと肝心の目標を達成できたかどうかの評価が定まらないことになる。

③設定根拠が明確であること

→ 数字自体を設定するときに、問題の発生件数ゼロとか、認知度100%とかで設定できるときがあるかも知れないが多くの場合は過年度こうであるとか、今までこのように推移してきたので将来はこうであろうとか、関連するデータとか既存の計画を踏まえて、根拠はこうなのでこうだと示さなければならない。根拠が明確でなければ達成できた、できなかったかを分析するやり方がないということになる。

□目標値の設定例

○「過去5年間の実績値の平均±10%以内」

→ 現状維持のようなものでもはっきり決めることで判断が明確になる。なぜこうなのかの根拠がなくてはいけないが、非常に明確な書き方

×「前年並みの水準を維持する」

→ こういう表現も目にするが、実際にその通りだと言った場合に、前年並みとはどれぐらいなのか、その場で適当に判断するというのはおかしいので、ここに書かれていないことが別に決まっていればそれはそれでいいが、それならそれで書いておいて欲しい。

×目標年度が平成62年度(基準値も途中実績値もなし)

→ 実際に政府の政策評価制度に出ていた例であるが、目標年度の62年度にいくらと

いうことで書いてある。この目標値が設定されたのはかなり前になるが、目標値を設定した時点で基準値はどうかと言う情報もないし、途中段階の実績値もなくて、平成62年度にこういう数字ですということだけ書かれてある。これではこの目標は何なのか実現する気があるのかということになる。

△「・・・の取組の結果・・・を実現」、×「・・・の取組を実施」

→ 文言で、何かを実現しますと書かれている。この場合は、目標年度にこれが出来ているかどうかの判断はできる。

→ これがなくて、・・・の取組の実施とだけ書かれている。この場合は達成できたかどうかは担当者の自己判断となり、評価とは言えない。×となる。

○/×過去3年間の伸び率が続くものとして設定

○/×10年後の目標値をもとに各年の目標値を等間隔に按分して設定

→ 両方とも明確な理由があれば○であるし、単に機械的に続くという話であれば×になる。もう一つ重要なことは、両方とも同じペースが持続するように見えるがこの2つは違うものである。グラフをイメージしてもらえばいいが、3年間の伸び率が続くというのは、グラフに書くと曲線的に伸びていくことになる。直線にはならない。経済成長と一緒にある。一方10年後の目標値については、現在値があり、各年の目標値を等間隔に、あるいは中間値を置いて、直線を引っ張って途中で割り振るということになる。こちらは同じ率ではなくて、同じ量だけ毎年伸びていくという設定なのでこの2つは根本的に意味が違う。この両者の区別があまりされていない。事業とか政策の性格によってどちらかということを考えなくてはいけない。

<演習>

Q3 アウトカム指標の目標値設定に対する考え方として望ましいものを以下から選んでください。

→ 目標値を設定するのではなく、目標設定の考え方が4つ書いてあるので、考え方として望ましいものはどれか。一つ一つ望ましいといえるかどうか○×△を考えてもらいたい。

(1)到達する可能性が小さいような高い水準は約束すべきでないから、目標値とすべきでない。

(2)一度設定した目標値は、短期間で変更すべきでない。

(3)目標値は、当該政策を所掌する部局が、保有する情報・データに基づき、責任をもって設定すべきである。

(4)基準値（直近の実績値）の10%増を目標値とする場合、政策によって10%増の意味は大きく異なる。

A3 → (4)だけが望ましく、外は何れも問題がある。

Q3(1)→なぜいけないのか。目標値には様々なスタンスがありうる。西尾先生の3分類を説明したが、諸条件に即したスタンスが望ましく、理想的な高い目標を掲げるべき政策もありうる。本当に実現できる目標を設定する考え方もあるが、できるだけ高いものを目指すという考え方は当然あってしかるべきであり、本来そうあるべきである。もちろん100本の政策について全て基本的に性格を統一したい、そのために実際に実現できるところに設定するというルールがあればそれは別である。個々の場面ではこのようになる。

Q3(2)→途中で目標を達成した場合や諸条件が変化した場合などは当然目標値を変更すべき。一旦決めてしまったものを変えることをしないというのは行政の内部事情

としてはあるかもしれないが、市民から見ればそれはおかしな話で、PDCA サイクルを実際に実効あるものにしようと思えば目標値の意味が変わったり、意味がなくなれば当然変えるべきである。行政の実際は決めたものをすぐに変えることには抵抗があるかもしれないが、評価論としては変えるべきはすぐに変えた方がよい。

Q3(3)→実施部局が、情報、経験とか現場を知っているのが最初に設定すべきであるが、評価担当部局によるチェック・調整や、意思決定者との交渉・合意などが望ましい。最近では自治体などは住民を加えて何らかの形で目標値を設定しているところもある。特にアウトカム指標などは住民の役割や事業者の役割があるものについてはそうした方々と一緒に目標を設定することもある。事業を知っている部局だけが設定すべきということは本来ではない。

Q3(4)→同じ 10%増でも、過年度実績の時系列の推移次第で、意味は大きく異なる。例えば過年度少しも伸びていないものについての 10%増と今まで過去 15%づつ伸びてきたものについての 10%増では意味が全然違ってくる。あるいは 10%という数字が大きく振れる可能性のある指標値の 10%増とあまり振れない 10%増では意味が違ってくる。数字の性格は様々であるので同じ 10%増でもいろんなケースがある。

(4) 測定（記録、計算）が信頼できること

- ・信頼性 (reliability) とは、同じ事象については、誰が何回測定しても同じ結果が得られること。

- 測定をするときには、きっちりと測定できる指標にしておかなければならない。妥当性が重要であるということは常にそうであるが、妥当性の次に重要なものとして信頼性ということもしばしば言われる。先ほどの落とし物を探すというときに明らかに探すとというのが信頼性の話なので、妥当性が大前提であるのは間違いない。

- ・信頼性を明らかにする（確認する）ためには「補助情報（データの履歴またはその所在情報）」が欠かせない。

- 信頼性で強調したいのは、この補助情報（データの履歴）がどこから来たのかを評価情報の中に示しておく、あるいは詳細はどこかにあるということや、データの根拠、出所はどこにあるのかを評価書の中で示しておくことは重要性である。

□指標値の信頼性に問題がある事例

×利用者満足度調査の結果、満足層が 94%。実は・・・

- 実は選択肢が 2 つしかなく、おおむね満足と不満の 2 つしか選択肢がない。ほとんどの人がおおむね満足を選んでいったという話であった。

×研修受講者のうち 75%が、修得した内容を実践（事後のアンケートによる）。実は・・・

- 本来受講すべき人がたくさんいるが、4 人しか出席できなかった。その人達はたまたまやる気があるのでいろんなハードルを超えて参加した人達で、そのうち 3 人が受講後実践しているということになると話はだいぶ変わってくることになる。

×市民の 60%が「ボランティア活動をしたい」と回答。実は・・・

- これもサンプルが少ない場合もあるかもしれないし、回収率が悪くて、意欲のある人だけが回答したものかもしれない。あるいは、アンケートで誘導してしまうこともある。例えば、ボランティアは意味のあることなので我が社は力を入れて取り組んでいる。ついてはあなたはどう思うかと聞いてしまうパターンがよくある。

▲合宿研修に参加した小学校の65%が、「〇〇ができるようになった」と回答（終了時のアンケート）

→ 小学校の合宿研修を行うと、なかなかそういう経験がないので、終了時のアンケートで高揚感や達成感があり、しっかりし、身の回りのことをよくするようになったという回答があるが、重要なのはもう少し後にするとどんな回答になるかということがある。

▲・・サービスの平均処理日数は5日間で、昨年から変化なし。

→ 時間が掛かるあるサービスを短縮してきたが、処理日数5日間でここにきて減らなくなって昨年と同じですといったときに、なぜもう少し細かい単位で作らないのかと聞けば、そもそもサービスの必要単位を1日単位でしか考えていないからということだったが、これは平均処理であるので元の単位は一日でも小数点に意味がある。例えば、昨年は5.1だったが、今年は5.2に悪化している場合もある。数字の信頼性の話である。

<演習>

Q4 ある市の評価書に次のような記載があったとする。測定の信頼性という観点から不足している（確認したい）情報にはどのようなものがあるでしょうか。考えられるものを列挙してください。

市民課の窓口サービスについては、今年度の顧客満足度調査の結果、満足層（「どちらかといえば満足している」以上の選択肢を選んだ回答者）の割合が50%であり、昨年度よりも5ポイント改善した。

→ 時間がないので後で解答も含めて、見ていてもらいたいが、市民課の窓口サービス調査をしたら昨年よりも満足したという回答が5ポイント増えたという発表があった。あるいは評価書にこのように書いてあった。この情報からはいろいろ分からないことがあるがそれがどれくらいあるかということです。解答を見る前に考えてから見てほしい。

A4 解答例

- ①「顧客満足度調査」の調査名、実施時期、集計結果・報告書等の所在（今年度・昨年度、以下も同様）
- ②調査客体のつかまえかた（標本抽出の方法など）、具体的な調査手順（実査の方法）、調査客体数（標本サイズ）、回収率（回答数）
- ③実際の質問・選択肢、調査票
- ④選択肢それぞれの選択割合
- ⑤回答者の偏りや構成比の変化の有無（性・年齢など）、及びその確認の有無
- ⑥誤差の見積もり、割合の変化は統計的に有意か

<演習>

Q5 数値の成果指標や目標値の設定が困難な場合、以下のような考え方は望ましいでしょうか。（○：望ましい、×：望ましくない、△：○と×の間。）

(1)ルール等で設定することが求められている場合、良い指標はなくても何とかして設定するのがよい。

(2)設定できない場合は、無理矢理に設定するよりも、空欄にしておく方がよい。

(3)数値による測定や設定が難しい場合、文言の指標設定、文言の目標設定もやむを得ないが、進捗状況の把握や目標達成状況の検証ができなければ意味がない。

(4)別の（成果指標や目標値によらない）方法による管理や評価を試みるべきである。

→ 今までの各論を総合しての話になる。数字の指標を設定できない、指標自体が考

えられない、指標は設定できても目標値が数字で設定できない、そのような場合にどうすればいいのか。ここに4つ考え方が書いてある。○か×か△かということで考えてもらいたい。時間がないので説明を聞きながら同時に考えてもらいたい。

A5(1)→ ルールで必ず成果指標を必ず設定しなさいと求められている。多くの場合でそのようになっている場合が多いと思うが、良い指標はないが、必ず設定しなさいとなっているので頑張って、少し無理をしてでも成果指標に近いものを設定しておくということについての質問です。

→ ×である。「良くない指標」に基づいてPDCAサイクルを回したらどうなるのか。指標の実績値はそこまで活用していないから大丈夫ですということ現場で聞くこともあるが、これでは何のために評価をしているのかということになる。だから、きちんとした指標を使わなければいけないし、きちんとした指標が作れないならば作らない方がよいことになる。でもルール上成果指標を作ることになっている場合、どう折り合いを付けるかは現場で判断をしなければならぬが本来はこういうことである。

A5(2)→ ×である。空欄は評価の放棄である。国民や市民から見てそうであるし、空欄を許すということは評価制度全体の品質を損なうことになる。何々の理由で設定しませんという理由が何処かになければおかしい。

A5(3)→ ○である。文言の指標であっても進捗状況の把握や目標達成の有無・程度が判断できなければいけない。それができなくてどのようにPDCAサイクルを回すのかということになる。

A5(4)→ △である。文言の指標で、目標管理、進行管理は限界、制約があるので、もう少し別の仕組みで、PDCAサイクルを回すにしても達成状況の把握をするようにする。別の制度を作るまでにいかないまでも、そういう事業は別の観点からチェックすることが必要ということになる。

おわりに

・数値の指標や目標値は手段であって目的ではない。本末転倒でヘンな数字を使えば評価の失墜・形骸化をもたらしかねない。(もちろん「出来る限り数値で・・・」が大前提

→ 数字の指標は目的ではなく道具にすぎないので、きちんとした道具を使うべきなのだけれども本末転倒でヘンな数字を使ってしまうと被害が甚大になる可能性があるので気をつけなくてはならない。

・「数をみせろ」と言われるから「数値目標流行」(某教科書に記述より)では困る！！

→ ある行政学者の教科書に数を見せろと世の中言われているので政府でも数値目標が流行しているのだと言っている部分があった。立派な行政学者がこのような指摘を教科書の中に書いてしまう状態とは、日本の公共部門の困った事態であり、行政学者にそうしたことを言われなければならないならなければいけない。

参考文献

(1)評価の客観性と統計数字の品質について

小野達也(2014)「政策評価の客観性と統計数字」『統計』65巻3号(2014年3月号)

小野達也(2014)「評価の客観性と統計数字の正しい活用のために一評価に用いられる統計数字の品質を巡って」『評価クォーターリー』No29

→ 数字の品質のような話であり、以前より強い関心、危機感を持っており、最近書いたものである。評価において数字でうそをつかないようにとの趣旨で書いたものである。

(2)指標の設定について

小野達也 (2009) 「業績測定型評価のかんどころ—プログラムのロジックと指標の妥当性—」『評価クォーターリー』No11

小野達也 (2015) 「評価のものさしを作る」『社会・政策の統計の見方と活用—データによる問題可決—』朝倉書店

(3)指標・目標値に基づく達成度評価について

小野達也 (2010) 「業績測定型評価のかんどころ(2)—目標設定・達成度評価における数・量の取扱い」『評価クォーターリー』No14

*小野達也 (2011) 「業績測定型評価における目標設定と達成度評価の妥当性—行政評価の形骸化を避けるための条件—」『地域学論集』(鳥取大学地域学部紀要) 8巻2号

→ 今日は指標を設定して、目標を設定するところまでの話をしたが目標を設定した後の話も重要である。実際に実績値が出て、達成度評価はどうするのかということもいろんなポイントがあるが指標設定も含めて、指標から始めて達成度評価をどう集約するのかということまで含めての話をここで書いている。

→ アスタリスクのついているのは、インターネットで入手可能なものである。一連の過程で何が起こり得るかということとあらゆる場合を想定して書いている。主として行政の職員の方に読んでいただくのを想定した論文である。

(4)府省と都道府県の現状

*小野達也 (2008) 「都道府県の評価の10年—定量評価のこれまでとこれから」『日本評価研究』8巻1号

*小野達也 (2013) 「政策評価と実績測定—府省の実績測定における計量・計数を巡って—」『日本評価研究』13巻2号

→ 日本評価学会の論文に書いたものである。全国の都道府県では数字をどれだけ正確に使っているのか、チェックした結果を論文にしたものである。

→ 次は政府の政策評価制度が目標管理型に移行しつつある段階の政府の評価制度と行政事業レビューも含めてその中で数字を適切に使っているのかをその時点で評価をしたものである。