

# 評価おぼえがき

## 財団の事業評価

笹川平和財団

理事長

入山 映

# 第1章

世のため、人のためになること。もっと具体的に言えば、少しでも住みやすい世の中にするためにはどうしたらよいのか、という事業を考えるのが財団の仕事だ。というのなら、何をどうやっても、それはそれで結構ではありませんか、ということになりそうなものだが、実はそうばかりも言っていられない事情がある。

まず第一が、よかれと思ったことが実は仇になった、というケースである。お節介な友人を持っている人が、時に経験するあれだと思っただけで差しつかえない。人道上の動機に基づいて援助物資を送り続けたら、受け手のほうがすっかり援助慣れしてしまって勤労意欲がなくなった、とか、甚だしきは地場の産業が壊滅してしまった、という例は作り話ではない。所得創出の目的で観光客誘致に力を入れたら、伝統文化が損なわれた。日本語教育をさかんにしようと思ったら、文化帝国主義だと非難された。零細企業に信用を供与するプログラムを提供したら、細々と続いていた相互扶助の「講」のシステムが破壊されてしまった。などというのはほんの一例だし、人権問題についてのアプローチもやりようによってはこの典型例を提供することになる。もちろん、ダム建設と水没する村、といった例のように、小の虫・大の虫の議論もあることだから、わずかなりともこの臭いがしたからといって避けて通ってばかりいる訳にもまいらない。しかし、動機がよいから、善意に基づくからといって結果が正当化されない場合があるのも当然だろう。実は最近の環境問題とか、生態系の破壊、といった場面でもこの点はよく議論される。アセスメントが有力な武器になることもあるが、その結論が必ずしも一義的でない場合、疑わしきは一切ダメ、という立場をとる人もいれば、有罪だと立証されないかぎり無罪だ、という人もいる。

第二は、資源配分の問題である。といったのでは何のことがわからないが、やらないよりはやったほうがよい、というだけでは、ある有限の資源をそれに振り向けるのに十分な条件とはいえない、ということだ。一個人のレベルで慈善事業に寄付したり、社会奉仕に参加したりするのならば、もちろんこむつかしい理屈は不要だが、社会機能の一端を担い、より多くの効果を期待するプロともなるとそうはいかない。税金を使う立場のお役所に、やらないよりはやったほうがよいから、といっておカネを使われたのではたまったものではないだろう。これには、「いまなぜそれなのか」という側面と、おカネを生かして使う工夫、という側面がある。前の方は、何らかの理由に基づいて作った優先順位（プライオリティ）に従って、やりたいことの仕訳をするという作業を意味し、後の方は、ある思いなり、目的なりをいかに効率良く実現するか、という方法論の問題である。この後の方の問題は別の厄介な問題を内包しているのだが、それについては後に触れることにする。いずれにせよ、ここで述べられている問題は、一般の企業や家計ではごく当然の話で、ことさらに議論するまでもないのだが、何か「よいこと」をする、という

動機に溺れると、こうした側面に無頓着になったり、甚だしい場合にはこうした考慮をはらうことが行為の純粋さを損なう、という見方まであったりする。

以上の二つが、「よいこと」だと思ったからといって、それを漫然とやっている訳にはいかない、という事情の主なものなのだが、この中から、民間非営利団体、あるいは財団にとって最も基本的な問題点が浮かび上がってくる。

それは他でもない業績評価の問題である。つまり、先に述べた二点で暗黙の前提とされているのは、仕事にはなんらかの基準で測定しうる成果あるいは結果があり、しかも、質の良い仕事、できのよい仕事、望ましい仕事と、そうではないものがある、という漠然とした見解であった。何故「漠然とした」とあえて但し書きを入れたかということ、そうした成果なり、質の高さを評価し判断する基準が相当に曖昧なものだからである。これは民間非営利団体、特に助成財団の仕事、一般企業と比較してみるとよりはっきりする。財団も企業も、ともに社会の、あるいは需要家の期待に応えようとする。その過程においてはさまざまな思い込みもあろうし、工夫もあろう。しかし企業の場合には、最終的には売れたか、売れなかったか、利益がでたか、赤字になったかという尺度が存在する。消費者が俗悪なものしか好まないかもしれない。より性能が劣る製品の方が売れたりするかもしれない。しかし市場によるテストはほとんどの場合において絶対的なものとして受け入れられ、確立している。特に費用対効果という意味でそれは判然とした尺度であるといつてよいだろう。それに比して、ある財団が、ある研究者の学会出席の旅費を助成したとしよう。それがどれほど意味があることだったか、他の研究者のほうがよくなかったか、他の学会ではどうだったか、そもそも旅費の助成と研究費の助成とどっちがよかったか。とても決定的な尺度が存在するようには思われない。

そうすると、「だってそれは良いことだと思ったからやったんです」というのとどこが違うのか、ということになるうし、それよりも、勝手な思い込みで好きなことをするのに、何故、その原資に対して税制面で優遇する必要があるんだ、<sup>※</sup>という議論を誘発しないか。役所が税金を使って何かをするのは、選挙を通じて選ばれた（従って民意を反映した）政治家、彼らによって策定された法律による施策、という理屈が一本通っている。一私人が、一民間団体が、好きなことをするのは訳が違う。全部とはいわないまでも、一部は税制面での優遇による原資（従って考えようによっては、本来は税金の型で国に入るべきであった原資）を使って仕事をするのならば、少なくともその仕事の出来不出来を評価し、責任をとらせる必要があるのではないか、という議論である。その基準を作るのがそんなに難しいのならば、日頃から役所が目を光らせましょう、管理監督を厳重に行ないましょう。そうでもしなければ国民の付託に応えられません、という話にもなりかねない。

※ 財団法人（公益法人）に対しては、法人の所得、法人に対する寄付、の両者について税政上の優遇措置がある。

## 第2章

財団の行う事業の質が問われている。それを判定する客観的基準があるのだろうか。もしかしてそれが難しければ、財団はその仕事の出来ばえについて、監督官庁の審査を受けなければならないのだろうか。前回に続いて、評価に伴う問題点を概観する。

それでは困る、というので何がしか客観的な評価基準を作ろうとしたとしよう。いったい何がそんなに難しいのだろうか。

評価と一言にいても、三十通りもあるそうで、一番古典的な定義は、事業を行うに当たって抱いていた目的・意図について、振り返ってみてそれが意味あるものであったか、を問い直すと同時に、それがいかに実現されたかという出来栄えに対する判断である、とされる。<sup>※)</sup>となると、出発点でいかに明確に目的を定義するかで勝負が決まる。事業実施者が評価を億劫がる要因のひとつはここにある。つまり、実施者がしなければいけないことは三つある訳で、一つは意図するところ、目的とするところに意味があるんだ、と言う立証。二つには、その意図を実現するために選んだ方法が妥当なものだ、あるいは優れているんだ、という論証。三つ目には事業が終わった後で（または行っている最中に）目的が達成できたかどうかを判断しやすいように、あるいは判断できるようなかたちでその目的を表現すること、である。評価技術がここでの論点ではないから細かくは立ち入らないが、特にこの三番目が問題になる。何故かといえば、一番目は実施者たるもの、どのみちはっきりさせねばことは始まらないのみならず、思い込みもあるから力説するにやぶさかではない（それでも大問題があるのだが、それについては後に触れる）。二番目もまず当然だといってよかろう。大西洋を横断するのに高い金を払ってコンコルドに乗りたければ、その必要性は説かねばなるまい。三番目となると話は違う。何らかの基準が指標を用いて判断が計量ができるように、といった途端に実施者は神の前に跪くような気がしてしまうのだ。本当は信じていないかもしれないのに、である。

例えば、あまり「日本では知られていない考え方を紹介する」講演会を企画したとしよう。括弧の中が事業目的だが、もちろんそれだけでは甘い、ということになる。どんな人々に、何人くらい来てもらいたい？ 本当ならば聞いてほしい人がいないではないんだが、来てくれるかな？ 出席率が悪ければ事業評価が低いのかな？ それなら目標値は控えめにしておくか。マスコミは取り上げてくれるのだろうか、その目安はどうやって決めればいいのか。講演はどれくらい理解されれば成功なのか、さてよ、理解度調査にはアンケート集計が必要だな、その費用は？ 理解度を示す尺度って何だろう？ ちょっと待った。いったいこの講演会は、こんな瑣末なことで「評価」されるのだろうか。と、まあこんな具合である。要するに、こんな神に跪いたり、帰依したつもりはない、という言い分であって、企業人がバランス・シートや利益率に跪くのは、そもそも納得づくでやっている。それとは訳が違うんで、私の志はそこにはありませんということになる。もっとも、これははなにも財団や「よいこと」をやる人に限った話ではない。自分の教育の成果を、有名大学への入学者数や共通テストの点数で測られたのでは不満な高校教師も多か

ろう。

もちろんこれは相当に粗雑な議論である。何よりも、事業成果の効率評価、すなわち formative evaluation と、そもそも論ともいべき summative evaluation との混同はたやすい。のみならず、評価基準の客観性という見地からは、何も議論されてないに等しい。つまり、こういうものが「質の高い」ものだとして世の中に受け入れられた、あるいは納得された、というものがあるかどうか、という疑問点に正面きって答えることになっていない。にもかかわらずこうした愚痴めいた話に本編にご登場頂いた訳は、その中に、なかなか本質をついた部分が垣間見えるからに他ならない。それは二つの側面において顕著である。

まず第一には、価値とその測定手段の混同である。オリンピックの百メートル競走ではコンマ・ゼロ一秒まで「客観的に」測定して金メダルが決まる。しかし、誰もコンマ・ゼロ一秒早く走ることの意味や価値は問わない。もっと極端な例を挙げれば、青少年に有害な製品を作っても、儲かってさえいれば利益指標はクリアしている。ところが世のため人のため、という次元になると、「志」が結構ならばそれ以上は問わないでよいのか、それとも一応そここのところも疑ってかかる代わり、思惑どおり結構なことだったらそれで終わり、ということになるのか。あるいは「その」志は与件と認識して、それが「どう」巧みに実行に移されたかを問題にして、それを測定しようというのか。それとも逆に、巧みに実行されてさえいれば、「志」の方は問わない(!)のか。あるいはその全部なのか。事態混沌として、いった何が問題になっているのか、論点が整理されていない。第二に競争とそれを提供する市場の存在の有無という点である。先に触れた資源配分の問題として考えたときに、ある「世のため・人のため」の行為（または商品・サービス）には類似のそれが存在して、あれよりはこれ、という選択の可能性があるのか、それはどのように提供されているのか。それとも、あれかこれかの選択的提供が問題なのではなく、提供されたものがある一定の水準に達しているかどうかだけが問題なのか、という点である。さらに、レベルが低いサービスしか提供できないのならば資源を他に「振り向ける」として、そのメカニズムがあるかどうか。さらに、選択といい、振り向けるという、それは誰がするのか。つまり需要と供給が会う市場のようなものは存在しているのか、その際の媒介パラメーターは何か。

唐突に市場の話が出てきて戸惑われるかもしれないが、実はこの問題は評価一般と裏腹の関係にある。つまり、仮にある事業の評価結果が劣っていたとして、それは事業の差し替えを意味するのか、それとも何らかの形での「その」事業の改善しか方法がないのか、という評価結果の効用についての大きな問題を提起することになる訳だ。

この二つの問題は、実は民間非営利団体の活動の評価をめぐる核心部分なので、少し突っ込んだ議論してみる必要がある。先に触れた。「意図」、「方法」、「結果」という三つのステップについてこのあたりを調べてみたい。

<sup>※)</sup> Michael Quinn Patton, *Utilization-Focused Evaluation* (Beverly Hills: Sage, 1988) によれば、評価には事業が意味があったかどうかを総合的に判断するもの (summative evaluation) と、事業の開発・改善に役立つ情報を集めるもの (formative evaluation) があるという。

## 第3章

第2章までに評価をめぐるいくつかの問題点が明らかにされた。これから数回にわたって、プロジェクトの選定・実施に当たって、「評価」では具体的に何がどう問題になるのかを調べてみる。

・「意図」の評価（その1）

まず、ある「よき」意図をもって何かをしようとした。

そこで先ず問われるのは、先にも見たように「よき」意図にも関わらず

その意図が結果として仇にならない保証

であり、それと同時に、

その意図が世に求められているものであるか否かの検証

つまり、ひとりよがりではなく客観性があることの証明である。さらに、

それよりもっと求められているものはないか。

世の中のプライオリティから見てどうだ、という吟味が必要だと考える人もいるかもしれない。

このうち、はアセスメントの問題である。どこまでの深読みを求めるか、またどれほどの労力を費やすか、というも議論であるが、それよりも黒白両論併記の結論になった際に、ゴーサインを出すか出さぬかが問題だろう。結論から言ってしまうと、これには「かくあるべき」という規範がある訳ではなく、個々の行動主体の選考と主義に依存する。その意味では、この評価プロセスは客観的基準を提供するものではなく、むしろこの読み込みをしておくことによって、望ましからざる副産物や意図せざる結果を最小限に留めうる、という実践的な効果が主たる機能である。ここでよく議論されるのが、アセスメントを行ったにも関わらず、望ましからざる結果を生じてしまうケースである。これを、一般企業の商品開発における事前の市場調査・販売予測と対比して、判断を誤った結果赤字になってしまったのと同じであるとする。そうすると、部門別経理などによって、ある商品開発計画の赤字・黒字がはっきりするほどに、「よいこと」をする場合の意図と結果との乖離を測定・判定する手段が存在しているのか、が問われることになる。ついでにいえば、予想とは違ったが結果はそれなりだったというケース（いささか古い例だが、名神高速道が開通してみたら、貨物自動車と旅客自動車の比率が最初の予想とは全く逆だったのは一例）をどう評価するか、という面白い話題もあるのだが、これらの点は後の の中心課題だから、そちらに譲ってここでは詳説しない。

他方 と については、世の中でそれが求められているかどうかという事実なり、その度合いなりをどうやって測るのか、という極めて具体的な問題（客観的基準）が中心になるが、時間軸がこれに加わると話は一層ややこしくなる。つまり、「いま」求められていることだけ行ってよいのか。民間非営利団体にとっては、先駆性こそが大事なことではないのか。何でそんなことにこだわるんだ、という意見である。実はこの他に、市場性の問題も後ろに隠れている。すなわち、どうも世の中ではそれほど求められていないらしい、

とか、もっと大事なテーマがあるらしい、という話になった時に、それでは他のテーマに置き換えられるものだろうか、どうしたら乗り換えられるのか、という点である。

ここで原点に立ち返ってみよう。そもそもあなたにはやりたい、やろうと思っていることがある。その意図（と実際に行ったこと）が客観的にみて意味のあることなのか、また質的にみて優れたものかどうかを検証する方法があるだろうか、というところから議論は始まっている。そしてここでは「意図」に着目し、それが現在の社会において求められているものかどうかを確認する作業から始めようとしている。いうまでもなく、これは一方でああなたの「意図」が社会的な支持を受けていることを検証しようとする訳で、この試みに成功すれば（ということはそのための作業方法が確定したうえで、だが）先ず第一のステップはクリアしたことになる。ところが、その試みに成功しなかったらどうなるか。

不成功には二通りあって、作業の結果が思わしくない、すなわち、どうも今の社会の要請とは言い難いようですよ、となるか、あるいはそもそも「客観的」な作業方法が見つからないか、なのだが、まず最初のケースでは、あなたの「意図」はおよそ行意味・価値がない、ということになるのだろうか。

ここで先の時間軸が登場する。民間非営利団体の存在意義の一つはその活動の先見性であり、時代を先取りして行くことにあるとすれば、やろうとしていることに「いま」社会の支持がなく、理解がないからといって、それがネガティブな評価につながるのだろうか、という疑問である。将来生まれてくるかもしれない需要を科学的、客観的に予測して把握するなどというのはできない相談だから、この議論は要するに「いま」の社会的要請にこだわることに意味がない、あるいは「いま」の社会要請は、あなたの意図に意味があるという証明にとって、十分条件ではあるが必要条件ではない、というに等しい。いうまでもなく、これは相当危ない議論であって、どんなおめでたい事業でも、この論理を駆使すればオールセーフになる可能性はある。そんなに野放しでは困るというのなら、先見性の名のもとであっても、最低限これだけは押さえておきたいという条件は何だろうか。一般企業の場合ならことは簡単で、儲かるかどうかかわからない夢物語に使える予算はこれだけです、と決めておけばよい。その判断が正しかったか無駄遣いだったかは神のみぞ知る、とは言わないまでも、棺を覆って定まる体のものでありまして、速断の限りではない。民間非営利団体の場合にもこの手が使えない訳ではないが、極端な話、組織を挙げてこの種の試みに没頭する、ということだったらどうなるか。

あたりさわりがあるかもしれないが、そこはお許しいただくとすれば、荒唐無稽としか認識されていないようなテーマ（永久機関、念力による物体移動、あるいは祈祷による疾病治療などを想定されたい）から、実現可能性についていささか夢物語の感の強いもの（直接民主制実現のためのコンピュータ・システム、世界連邦における憲法案）まで、これに大真面目で取り組みたいとするものを、これはよいが、あれはだめ、という線がいったい引けるものだろうか、というのが極端な設問である。

## 第4章

第3章では、具体的な評価作業の手始めとして「意図」の評価を取り上げた。そこで問題になったのが、将来を見越した先見性のあるプログラムに対して、「いま」の時点でその適否を論ずることができるものだろうか、という点であった。

### 「意図」の評価（その2）

そんなことができる訳がない、というのが一つの立場だろう。だからどうなんだ、というところで二つの相反する結論の出し方がある。

一つは、だから、「いま」の社会の要請を論ずるときに先見性や存在意義を引き合いに出すのはやめた方がよい、とするもの。二つめは逆に、それがいかなるものであれ、（先見性があると称する）意図は無差別に認めるしかない、とするものだろう。第一の結論は再び先の議論、すなわち「いま」の社会的要請をいかに把握するか、に戻ることになるが、二番目の方は、そもそも意図に対する評価は意味がない、ということになりそう。そのかぎりにおいて、この立場の支持者はここでの議論を打ち切ってのステップに進むことになる。ただし、どんなに脳天気な意図でも無条件で認めて大丈夫なのか、という懸念は依然として残る。良識の正規分布と淘汰の必然性がこの懸念に対する答えだろうと思われる。つまり、どんな意図であれ（公序良俗に反したりしないかぎり）民主主義社会においては市民権を持つ。ある程度の社会的な納得さえ得られないような意図であっても、もちろん市民権はある。しかし、大数観察をすれば、おそらくそうした意図の存在、実現しようとする人々の数、は社会において応分の割合しか占めないだろうし、中長期で見れば、余りにも突飛なアイデアやテーマで、実現性の乏しいものは、それを支持したり、携わろうとする人々の数が漸減してゆく。逆にいかに少数の人にしか理解されていないか、筋目正しい話というのはいずれ支持者が増えてゆく。こうして淘汰されていくだろう、ということだ。そしてこの観測を歴史的に実証することはそんなに難しくはないと思われる。ただし、すべての歴史的検証がそうであるように、どれほどの時間の長さで捉えるか、によって結論が逆になる可能性は常に存在している。だから、与えられた時間枠の中で、どれほど実際上の意味がある議論なのかについて疑問がない訳ではない。にもかかわらずここで「大数の法則」まがいの議論を持ち出したのはいささか訳がある。

それは、この議論が実は民間非営利組織の存在意義に言及することに連なるからに他ならない。そもそも世の大多数によって現に支持され、必要とされているが、市場原理にはうまく適応しない需要に対処するのは政府の仕事である。政府はそれを自ら供給しようとするかもしれないし、それが得策でなかったり、お役所がやるより誰かほかの人がやったほうがはるかにうまくいくのなら、お金だけは税金を財源とするが、民間非営利組織にやらせることもあるだろう（国家財政がそれを許さない場合には、本来国のやるべきことをそれ以外の主体に委ねることもあるかもしれない。しかしそれは

あくまでも例外的・緊急避難的な事象である）。少数意見とか、世の大方の支持を得るには至っていないが多くの可能性を秘めた創意を育む、という仕事は明らかに政府が行うのには適していないが、それはその存在の重要性を否定するものではない。それどころか、ある社会がそうした立場をどれほど尊重し、存在することを奨励しているか、は社会の成熟度と民主度のバロメーターであるといつてよい。そうしたありようを淘汰し、進化させるのは市民社会そのものであって政府ではない。この部分に対する資源配分が過度にわたって、最大多数の最大幸福にとって明らかに阻害要因になる（それを決めるのも市民であるが）ことが証明されるまでは、最大限の配慮がなされるべきだ、という理由づけをここに見い出すことができる。さて、再び先見性の議論に戻る。

目利きはできる、という人もいるだろう。本物といい加減なものを見分けるのがプロフェッショナルリズムだ、という立場である。しかるべきプロが財団の中（あるいはそのまわり）にいて、きっちり判断すれば、先見性の名のもとにとんでもないものが紛れ込むことは防止できる、ということだ。そもそもプロというのは、ある分野の知識や技量によってお金を稼いでいる、という意味が一般的だが、要するに素人よりは一芸の上手である。プロの暮打ちはアマより強いし、プロゴルファーも同様。それでは財団のプロとはいったい何だろう。何が素人よりもうまくできるというのだろうか。事業型であれ、助成型であれ、ある専門分野の研究とか、具体的な学問領域に特化している場合には話は割合簡単だが、学際領域のテーマとか、社会開発とか国際交流のような分野を扱っている場合にはいささか悩ましいものがある。というのは、そうした分野で仕事をするプロというのは、高度の専門知識の蓄積によって、「どうだ参ったか」というのにもピッタリしないし、水際立った技量の冴えが衆目に明らかだ、ということでもない。それではどんなイメージだろうか、という関心をひくのだが、財団プロ資質論は当面本稿の主題ではないからこれ以上立ち入らない。ただその存在の必要性については、ここでの「意図」をめぐってのそれ以外にも、財団の活動全般についてかなり汎用性のある論点だ、ということ指摘しておくにとどめる。

さて、プロだからといって万能な訳でも何でも無いのは、プロ野球の選手が10回に3回ヒットを打てば上出来だというあたりの事情に明らかで、それどころかプロ中のプロが目利きをしているはずの博物館がニセものをつかまされたりすることもあるのは周知のとおり。だから、例え結果が思わしくなくても「あの人がそう言ったんなら」しょうがない、と納得するかどうかだということになったりしがちである。それがあらぬか、高名な選考委員を並べたり、専門委員に名士を連ねることでこの要件をクリアしようとするのも決して稀ではない。しかしここでの論点がそんなところにはないのは自明であって、要は客観性の基準がプロ（と認められる人々）の存在によって置き換える可能性の示唆が主題であった。これはかなり重要な点であって、評価全般について興味深い解決策を提示する可能性があるが、ここでは意図の先見性についての検証に留める。

## 第5章

「意図」の評価をめぐって、おおよその可能性が明らかになった。ただひとつ残っているのが、「よいこと」相互間の互換性・つまり、あれよりはこれをやってみよう、という選択がどのように可能であるか、という問題である。

### 「意図」の評価（その3）

こうして問題は再び、あなたの意図が社会的支持を受けているかどうかを確かめる客観的な作業方法に戻るようになった。社会的支持が存在するかどうかを定性的に検証することは、それほど難しいことではない。むしろ問題とされるのは、支持の度合、つまり「これ」と「あれ」を比べたときに、どちらがどれほど強いかわかると比較することに意味があるか。すなわち望むべくば、より多くの人々に望まれ支持されていることを優先すべきなのか、そうでないか。そもそも複数の「意図」相互間に比較計量する手段があるのか。何らかの理由で当初のあなたの「意図」を、より優れた、あるいはより多くの人々に望まれるものに変更しようとした場合、それは可能なのか、といった点になる。

十万人に一人の難病の治療法を研究するよりは、水虫の特効薬を開発するほうがはるかに多くの人々のためになる、だから難病はやめておけ、とは誰も言わないだろう。さりとて無数に存在する社会的要請を定量化して順序をつける、などというのは、計画経済ではあるまいし可能だと思う人が存在するとは考えにくい。そうするとここでの問題は、意図といい、社会的支持という、そうしたものの相互間に質の差、あるいはレベルの違いを見い出すことが可能だろうか、という点に限ってよいのではないと思われる。つまり、ただ単にあなたが「やりたい」と思ったから、あるいは「よいことだ」と考えたから、というだけで満足するのではなく、ひとつには「その」あなたの意図が、質的なレベルがどの程度のものであるかを判断する手法があるのか、また、その判断の結果、より質の高いものを志向するほうがよい、と考えた時にどうやって新しいものを見つけてことができるのか、ということである。

このうち前段は、先に先見性について議論した際の選択肢が援用できることから再度詳説はしない。ここでは後段、すなわち「よいこと」相互間の競争と市場性について少し調べておくことにしたい。

話は二つに分かれ、ひとつは意図そのものの選択に関わり、二つ目は意図の実施主体に関わる。いずれも理論的な可能性と実際上のそれとのすりあわせである。

第一の点は、さまざまなレベルで、そしてさまざまな機会に発生する。一番解りやすい例は、芸術支援をしようとした際に、音楽にするのか、美術にするか、それとも演劇、文学、どれをとるか。仮に音楽をとったとして、古典か現代か、はたまた邦楽か西洋音楽か。作曲家か演奏家か、楽器制作者か、それとも音楽史研究家か。あるいはそもそも現在の社会において、芸術支援を考えるのと、環境問題、途上国での開発問題、核軍縮などの問題にエネルギーを注ぐ

のとでは、どちらに意味があるのか。これまでに見たように、専門知識の助言を得たり、アセスメントや社会性についての検証を試みることによって、当初あなたが抱いた意図について最低限（あるいは最大限）の質的なチェックは可能だという立場に立とう。その結果が仮にあまり満足のいくものではなかったとして、それに代わるもの、おおげさにいえば森羅万象を網羅しかねない代替可能性と、あなたの限られた情報源を結び付けることはできる相談なのだろうか、ということである。そして第二には、なんらかの意図が定まったとして、それを実現してくれる人や組織をどうやって見つけるのか、ということになる。このこと自体は「評価」とあまり関係がないように思われるかもしれない。しかし実はそうではなく、質の高い成果や成果の質を高めて行く努力にとって、この要素は不可欠なものである。

需要と供給が出会う場所、あるいは仕組みが市場である。市場の大きさと質が需要・供給の原因となり結果となって、高度の市場経済、あるいは貨幣経済を発達させてきたことは周知の事実であり、この仕組みの発展がないところでは満たされぬ需要、起こるべくして起こらない供給が発生するのは、最近までの計画経済の歩みを一瞥すれば足りる。そして歴史的に見れば、この市場制度の確立を自然発生と自律的な成長に俟つだけではなく、その成熟を下支えするために各種の経済法、取引法などさまざまな法的措置が講じられてきたことが知られるであろう。ここでいう法的規制が、規制緩和の対象と見なされているそれではないことは明らかである。ところが、ここでわれわれが議論している財貨・サービス（「世のため、人のため」のそれ。通常公共財と呼ばれる）については、もともと市場性になじまないという原理的な性格からこうしたインフラ要件の整備が非常に遅れている。政府情報の公開の遅れはその代表的な例であり、税制を含む各種制度も、この市場の成熟に向けてというよりは、むしろ逆の方向に機能していることもよく指摘される。<sup>注1)</sup>市場は競争を生み、競争は質の向上をもたらす。その意味において、需要と供給の出会う場所の社会装置化、<sup>注2)</sup>それによる競争を通じての質の向上は、個々の主体による評価の試みと共に重要な前提条件であり、いわば車の両輪であるといつてよい。<sup>注3)</sup>従って、よき「意図」の評価をめぐってより質の高い意図を模索する試みの一部は、同時にこうした仕組みの開発に振り向けられることが望ましい。またそれと同時に、民間非営利団体の仕事の出来不出来を評価し、管理監督しようとする努力よりは、市場条件を整備し、活動を成熟させるための要件を満たしていくほうがはるかに生産的であることもいうまでもない。

評価は、もちろんそれ自体が自己目的なのではなく、その結果によって改善を指向したり、場合によっては対象とされたことからの置き換え作業を行うことに意味がある。「よき」意図をめぐって、いかにそれが可能であるかの検討は、その二つを念頭において進めてきた訳だが、ひとまずこれくらいに止めて先に進むことにする。

<sup>注1)</sup> 雨宮孝子『フィランソロピーと社会』ダイヤモンド社291頁以下。また、総合研究開発機構『市民公益活動基盤整備に関する調査研究』149頁以下。

<sup>注2)</sup> 拙著『社会現象としての財団』NHK出版178頁以下を参照。

<sup>注3)</sup> 仮にこちらの条件が整備されていないと、評価の機能はある事業の実行方法の改善、ということにしかならない。「他の」事業を探したり、それに振り替えたりすることが困難だからである。

## 第6章

事業評価について、第一段階である「意図」の評価に関する可能性を調べてきた。引き続いて、事業の方法論と、事業成果を振り返っての評価に話は進むことになる。

### ・ 方法論の評価（その1）

意図についての評価が一段落すると、次に、何かをする、その方法論が決定される。

そこでは、やり方・手口の巧拙が問題とされる訳だが、それには前提条件として、巧拙を判断する基準（の有無）が問われよう。

先に見た「意図」とこの「方法論」は表裏一体になっていること、あるいは区別して論じることに意味がないような場合も少なくない。例えば褒賞（ナントカ賞授賞の類）とか奨学金の付与、あるいは旅費や出版に対する助成、さらには人道援助といわれるものの多くがそれである。しかしここで敢えて観念的に区分して議論しようとするのは、もっぱら操作的（operational）な理由による。つまり、方法論の巧拙を問うというのは、もっとうまいやり方はないのか、なかったのかという意味であり、巧拙を判断する基準を、方法論改善のために必要な情報の種類、と読み替えてみると、意図と手段を区分して見る実益がはっきりする。評価それ自体は、いまだ存在していない新たなものを創造する行為ではない。したがって、巧みさを判定する客観的基準（第2章の百メートル競走の例でいえば、コマ・ゼロ秒の測定）がかりに見い出されたとしても、それがより巧みな方法（スタートのタイミング、手の振り方）を考案することと同義ではないのは明らかだ。それならば原点に立ち返って、よりよい方法を見つけるために必要な情報収集が評価機能として浮上してくる。

観念の区分を実例で示せば、例えば日米貿易摩擦の険悪化を緩和するためには、相互理解の深度化が必要だと考えたとして。そして相互理解を推進すべく、両国の国会議員の交流・討論会を開いたとする。険悪化緩和は意図である。交流・討論会は手段だ。しかし相互理解深度化は意図だろうか、手段だろうか。この議論はそれ自体としてはさして実りのあるものとは思われない。日米貿易摩擦の険悪化が極度に進行していて、いまさら相互理解の深度化などという悠長なことをやられてはならない、もっと即効性のある手段を選ぶべきだ、というのと、どんなに一触即発であれ、相互理解の努力は続けるべきだ、と互いに言いつつっていたのではおおよそ議論にならない。しかし、この両者を評価の対象として区分することには明らかに実益がある。続けるべきだ、という方は意図として評価すればよいし、ほかの方法を選ぶべきだ、というのなら（険悪化防止のための）相互理解深度化にとって有用な情報を集めればよいことになるからだ。

有用な情報とは何か。ここでは事後に、つまり事業が終了した時点のことに限っている訳だが、当初期待した目的がいかに実現されたかの検証に役立つもの（これについては後の「成果についての評価」に譲る）の他に、「もう一度同じことをするんだったら、今度はこうやりたい」という知恵があるはずだ。それが求められてい

る情報である。ところでその知恵なるものは、当初計画の失敗によって得られたのかもしれないし、あるいは経験に学んで初めて得たものかもしれない。「評価」という表現にこだわるとこのあたりが気になる向きも出てこようし、人によっては、とんでもない馬鹿をしたのとそうでないものとを区別しなければ気がすまないかもしれない。しかし文脈から明らかのように、ここではそれを問うてはいない。財団として計画段階での杜撰さを奨励している訳でもなんでもないが、特に助成財団では失敗を怖れて安全第一に陥ることの弊害のほうを重く見るからである。この態度こそがこの組織を創意に満ち、時代を先取りするものにしていく要因の一つだといってもよい。しかし同時に、この立場が無責任とか安易のそりを受け易いのもまた事実で、先に見た評価の必要性の論拠、すなわち「勝手な思い込みで好きなことをする」のに対するチェックというのはこの延長線上にある考え方だといってもよい。そうすると、何のことはない「意図」の評価をめぐって、ひとりよがりのおめでたい「意図」をどうやってはじき出すか、という問題意識が形を変えて、とんでもない馬鹿をやったらどう責任をとらせるか、という問いかけになっていることが知られるであろう。

極めて根強い議論なので煩を厭わずにもう一度繰り返すとこういうことだ。馬鹿をやり、おめでたい見込みで仕事をしなさい。営利企業ならば黒字・赤字で結果が出る。赤字決算を三期も続ければ勝負あった。会社が破産しないまでも、経営者ならば退陣だし、セールスマンならクビか左遷だ、という訳である。お役所でいえば会計検査院とか、世論の指弾による掣肘だろうか。それに比べて財団ではどうですか、放漫経営したい放題、一体誰がどう責任をとるというのですか。それとも仕事の出来・不出来についての客観的評価基準でもあるというのですか、ここはやはり監督官庁の出番でしょう。

これに対する反論はこうだった。ちょっと待って下さい。仕事の出来・不出来をはかる物差しがあるかないかの話と、責任を取る取らないがごっちゃになっていませんか。物差しについていえば、世の中何でもヨコに比較できる訳ではないでしょう。価値のあるなしといったお話は、それなりの玄人に頼るしかない時もあるれば、歴史に聴く他はないことだってありはしませんか。物差しがないから出鱈目だというのは無茶な議論というものです。ついでに言わせて頂ければ、世の中に創意工夫が大事だと思うのなら、遊びと冒険抜きではあり得ない。よほど人様にご迷惑な破目のはずし方ならいざ知らず、なんでもかんでもたがをはめればよい、というのこそ困った考えというものでしょう。

どうも議論がかみあっているとは思えない。最初に投げかけられた質問が相当にいい加減なものだ、というのにも半分責任はあるが、一方客観的な評価基準は可能なかどうかについては、それは無理だ、と言っているようだ。では品質管理はどうするのですか、と聞かれると、（プロである）当事者にお任せ下さい、という答えしか返ってこない。それならば何で評価についてくどくどと書いているんだ、ということになる。

## 第7章

事業方法論の評価と関連して、再び、どうしようもなくお粗末な事業を見付け出すことこそ評価の主目的ではないか、という問いかけがなされる。もしかその答えがイエスで、しかもそれが不可能だという結論になるのなら、一体評価について論じているのは何故なのだろうか。

### 方法論の評価（その2）

もちろんそれにはそれなりの訳があるのでありまして、言いたいのは、「評価」の機能にはいろいろあって、その機能が何であるか合意するところからすべてが始まる。機能、すなわち評価の目的によっては手段も変わってくるし、角度もさまざまです、ということだ。「意図」について、「作業方法」について、あるいは財団の先見性、プロジェクトの市場性についてこれまで検討してきたのはその理由による。評価・測定が一義的なものではないのは何も財団に限ったことではない。百メートル競争で大事なものは順位なのか、記録なのか。ゴルフだって、マッチプレーかストロークプレーかによって結果は変わる。そういう立場で考えて見ると、いま議論している「作業方法」の評価には、将来に向けての事業執行方法改善という機能がある、といっている訳だ。ならば、新たな手法開発のための評価、さらには馬鹿げた事業を発見するための評価も当然存在しようというものだ。このうち、前のほうは「もう一度同じ目的で仕事をするんだったら、今度はこういう方法でやりたい」という意味での情報収集になる。これと「もう一度同じことをするんだったら、今度はこうやりたい」との違いは、方法は同じだが扱い方・やり方に工夫を加えるのか、方法を変えてしまうのか、とすることになる。前章の国会議員交流の例でいえば、討論会のテーマをいくつかに絞った上で、あらかじめ参加者に意見をまとめておくよう依頼しておけばよかった、とか、議員の選出選挙区の多様性に対する配慮がもう少しあったほうがよかった、というのがやり方の工夫ならば、公開討論会というフォーマットではあまり本音の議論ができないから、非公開の泊り込み合宿のようなもののほうがよい、とか、さらに、国政レベルの議員よりは、地方議員の方がよかった、というのが手法の変更である。

後のほうはどうだろうか。これまでのところは空回りをしているように見えた議論も、この延長線上で仕訳して見ると、三つに分けることができるのは明らかだろう。一つめはもちろん馬鹿げた事業とは何ぞやから始まる。稚拙、劣悪、有害（本稿冒頭の例を想起されたい）、何といってもよいが、要はその道の玄人三人が三人とも「これはひどい」という体のものだとしておこうか。そんなものが大手を振ってまかり通らないように予防するためにはどうするか。これは既にして問わず語りに明らかであって、しっかりした人に計画をたてさせなさい、その過程でしかるべき人の知恵を借りなさい、それでも漏れたら理事会だって評議員会だってあるでしょう、というルーティンの話になる。二つ目は、まさか馬鹿なことが起こってはいはしないだろうね、と確かめたい、あるいは起こっていません、と報告させたい。確かめたかったり、報告をききたがるのが誰でも

るかが問題なのだが、上記ルーティンがしっかり行われているかどうか理事会が、監事が知りたいというのならこれは当然の話である。一步進めて、社会に対する情報公開の義務として、このあたりをつまびらかにすべく法制化したい、というのも傾聴に値する議論だろう。<sup>※</sup>三つ目は、馬鹿げたことをやったら責任を問う、という話だが、組織内部でこの問題にどう取り組むか、というのなら格別、裁判官か風紀委員のような人がいて、かしこんでその方のご判断を承る、という感覚の議論ならばナンセンスであることはいうまでもない。

少し本題から離れるが、この「責任」問題について組織内部でどう取り組むか、というあたりについて触れておきたい。というのも、評価の目的が何であれ、それはより良いプログラムを作り上げて行きたい、というのがひとつの大きな動機なのだが、それと同時に、あるいはそれにも増して、成果の善し悪しを判定した上で因果応報、それなりの措置をとるべきではありませんか、という意見が結構少なくないからだ。これには二つの側面があって、一つは組織内部に先に見たルーティンが確立されているかどうか、ルーティン改善・改良に向けての努力が払われているか、というマネジメントの問題であり、いま一つは担当スタッフの練度向上をどう計るか、技量・達成度の評価を個人の待遇にどう反映させてゆくか、という人事管理上の問題である。

マネジメントについていえば、有能で注意深い経営陣には問題はないのだから、無能でいい加減なそれをどう排除するのか、という議論になる。スジ論は明らかで、経営陣を監督する理事会（および評議員会）がしっかりせよ、となるのだが、これは株主総会に取締役会の出来不出来を見届けるというのと同じ問題を抱えている。他にもない「べき」論（sollen）と「である」論（sein）の違いであって、ある「べく」期待されているからといってそう「ある」訳ではない、というに尽きる。だからといって、頼りにならない理事会の代わりにどこかの機関が目光らせる、というのは、株主総会の代わりに会計検査院にさせる、というかつての共産主義国家みたくなことができるはずもないのと同様である。現実的な解決策は、活動内容の公開と当事者の資質向上プログラムの充実だろう。それによって内・外双方から活動のレベルアップを図ることになる。速効薬の好きな向きには余りに迂遠に感じられるかもしれない。しかし、そもそも民主主義とは迂遠なものなのだ。

人事管理は評価と表裏一体の関係にある。これは一般企業の事情と同様であるが、ひとつ基本的に違うのは、民間非営利団体の組織原理はヒエラルキーによるものではなく、使命感の共有と参加の意識により重点をおくことである。従って、技量・達成度の評価が昇進・昇給に直結するというよりは、むしろプロフェSSIONナリズムに裏打ちされた達成感の充足という側面がより重視される。もちろん財団職員として糧を食べて生きている訳ではないが、ボランティア精神が民間非営利団体の基本的構成要素の一つであることは忘れられてはならない。

<sup>※</sup> 情報公開のやり方と内容はいろいろ考えられるが、ここでの問題に限れば、事業方法選定の当事者、理事・評議員・監事の氏名公開ということになる。ただし、ノーベル賞の選考に典型的であるように、判断に対する外部からの影響を排除する意味で選考委員の氏名を非公開とすることは十分にありうる。秘密主義と公開性の問題である。

## 第8章

これまでは、どちらかという評価という作業をとりまく背景、さらには考え方の基盤を中心に話が進められた。今回から、いよいよ最もポピュラーな「成果」評価が登場する。

### ・ 成果についての評価（その1）

「意図」の評価、「方法論」の評価、と説き進んできて、いまだ答えられないで残っているのは、期待された成果があがったかどうか、すなわち、いうところの「成果」についての評価だけになった。これは評価についての一番古典的な考え方であり、通常「事業評価」という場合に誰しもがまず念頭に置くものだといってよい。

この議論はこの小論の冒頭にも触れたように、「結果として何を期待するかの規定」から始まることになる。これは要するに、最初に期待した成果と後で実際に出た成果を照らし合わせようというのだから、当初目的の記述が漠としたものであれば、成果との対照も朦朧としたものになる。あるいは極端な場合、結果が出てから後付けの論理で「これが目的だった」とした<sup>※1)</sup>のでは何のための評価かわからない。だから、当初のそれをいかに綿密に規定するかが肝心で、そのためには目的の厳密な規定に没頭することになる、と考えられがちだ。そうすると、始まってもしない事業の成果を事細かに記述することにエネルギーを費やす、ということになって（第2章の講演会の例を想起されたい）実務家にはなはだ評判がよろしくない。のみならず、当然のことながらその後で

1. 結果がその期待に適合したか否かの判断が求められる訳で、そこでは
2. 判断の定量化の可能性、それに際しての基準の有無までが問題にされざるをえない。

以上は当初に予定した結果あるいは成果が、はたして目論見通りにいったかどうか、という検討である。しかし、水虫の薬にと思って開発したら毛生え薬に効いたらどうだろう。結果オーライではいけないのか、という疑問は残る。実はこれまで評価の試み、評価の歴史の中で、この結果オーライを含めていくつかの割り切れなさが痛感されてきた。それが先に触れた三十通りもある、といわれる評価方法になっているのだが、ここで、それらがどんなものを概観しておくことにする。<sup>※2)</sup> というのも、先にも述べたように、評価はある目的のもとになされるものであり、従ってその概観を通じて、人々がこれまでに何を意図し、どんな結果を求めたか、についておおよその知識を持つことになる。いく通りもある評価のパターンについての知識とは、とりもなおさず、場合に応じて最もピッタリした、つまり最も納得しやすい評価方法を模索した結果の一覧表になる。だからそれは同時に、プロジェクトによって（ある手法を前提とすれば）評価が簡単にできるものと、そうでないものがあることを示唆することになるだろう。これまでに述べてきたところから読者諸賢がお察しのように、この小論は、客観的な評価が難しそうな

場面、一義的な評価があまり期待できそうもない局面をことさらに選んで議論している。いわば現場の悩みについての「おぼえがき」な訳だ。だから一方では評価一般についての理論的考察でもなければ、網羅的なハウトゥーでもないことは、遅ればせながらお断りしておかなくてはならない。

さて、これまでに開発された評価の類型は、大別して目的に関するものと手法に関するものに二分される。

まず目的に関するものは、目的の多様性に応じて千差万別の観を呈するのだが、主なものを拾ってみると、先に触れた そのプログラムを続けたほうがよいか、やめてしまったほうがよいかを問う（summative evaluation）、そのプログラムを改善するためにはどんな方法があるかを問う（formative evaluation）という二つ。さらにここで論じている 目標がどの程度達成されたかを問うもののほか、

そのプログラムのもたらした現実の効果（意図したものであれ、なかれ）を問うものも可能である。評価結果をどう利用するかに着目すれば、許認可などのための最低限の基準を充たしているかどうか、ある時点である決断をするためにどんな情報が必要か、あるいは毛色が変わったところで、プロジェクトの成果周知のための評価、などなどの類型が考えられよう。新機材の投入や新製品の使用に当たっては、費用対効果評価、が可能な場合もあるだろうし、プログラム実施の当事者にとっては、努力度評価とか スタッフの人事評価も関心事だろう。需要家のニーズに着目すれば、ニーズ・アセスメントもこの範疇に入ってくる。

手法に関するものとしては、内部の人間（通常はプログラム・スタッフ）によって行うか、外部に頼むか。さらに 評価結果が数値化する定量的（quantitative）評価と、数値化が困難であるか、あるいはそれになじまない事業を対象とする叙述的・定性的（qualitative）評価、また、ある評価基準にてらして達成度を測定するものと、理由や効果は問わずに、いったいそのプロジェクト実施の結果何がおこったかだけをみようとするもの、などが主なものであろう。

こうして眺めてみると、評価の試みとは、何らかの目的意識の下でなされる他に、何らかの規範意識（あるレベルに達しているか）、あるいは制約要件（評価実施に際して動員できる手段は何か）の下に行われる場合が結構多い。逆に言うと、行おうとしている評価作業は何を目指したものであり、そのためにどのような物差しを用意しているのか、をはっきり認識してから事に当たらないかぎり無用の紛糾を生む、ということのようだ。そうすると、ここで試みようとしている成果評価についても、当初の意図をできるだけ詳細に記述して、後で成果と照らし合わせやすいようにする、という作業に一生懸命になる前に、そもそもなぜそんなことをするのか、なぜ結果と当初の意図との整合性が重要なのかをいま一度吟味しておくことから始めたほうがよさそうである。

※1) Pattonに言わせると、「先にダートに壁に刺しておいて、その周りに的を描くようなもの」だということになる。

※2) 以下は、M.Q. Patton, *Utilization-Focused Evaluation* (Beverly Hills: Sage, 1988)

## 第9章

当初の意図と結果をつきあわせてみる、というのが「成果」評価の原点だとすれば、そもそも何のためにつきあわせてみるのか、について考えた方がよい。余りにも自明かもしれないが、この点を曖昧にしておくと、後で議論が紛糾することになる。

### ・ 成果についての評価（その2）

結果がよかっただけで駄目なんで、それがそもそも意図され、計画されたような成功でなければならぬ、というのは一体なぜなのだろう。

一つには効率または反復可能性の問題であるのは明らかだ。つまり、思いもかけない成功（前章の水虫と毛生え薬の例を想起されたい）をおさめたからといって、もう一回思いがけない成功がおこる可能性は少ない。今回結構なことがおこったのを否定する理由もないが、その手口をお勧めする訳にはまいらない。「まちぼうけ」ではないが、切り株の前に座ってもう一回兎がころんでくれるのを期待するのは難しかろう、ということだ。

第二には、その裏腹だが、学習と教育効果の問題だ。ああなってしまうだろう、という目算あればこそ、どこが巧くてどこが拙かったかという話になる。どの前提が正しかったか、そうでなかったかも解ろうというものだ。さもないければ、成功の秘訣どころか失敗から学ぶことも出来かねる。

第三に成功（あるいは失敗）が、その企画をたてた人のおてがらなのか、それともただ運がよかった（あるいは悪かった）にすぎないのかをはっきりさせておかななくてはならない。ただし、この問題は上の一か二の理由に帰することができる部分が多い。運がよかった、その運が何で、その条件を除けば、という分析ができれば学習にもなるし、再度反復にあたっての指針にもなる。そうでない部分というのは、おてがらならば誉めてあげましょう、有能だと認めましょう、という側面にすぎない。

従って、最後の部分を別にすれば、「計画されたような」成功でなくてはならぬのは、学習効果の効率を上げる目的がある場合に限り、ということになる。これはしかし、何も財団だ、非営利だにかかわらず、世に組織といわれるものならばどこにだって当てはまる話であって、一般企業のマネジメントの心構えとほとんど選ぶところはない。

ここから導かれる議論は二つあって、一つは、それ以外の場合、すなわち学習効率を上げる以外の場合には結果だけを問えばよいということである。結果がある一定の基準を充たしているかどうかを問う、成果が何であるかをはっきりさせて、その意義を世に知らしめる、などという場合には、当初の意図がどこにあったかをいちいちつまびらかにする必要はない。苦役からは解放される訳だ。とはいえ、例えば助成財団における評価のメリットとして、財団自身の助成活動の改善、アカウンタビリティの証明、助成先への技

術支援、助成団体全体のノウハウ蓄積、組織活性化の5点が挙げられることがある<sup>※1)</sup>が、だとするとを除いてはすべてが「計画されたような」成功でなくてはならぬ場合だ、ということになる。苦役からの解放はそう簡単ではない。

もう一つは、効率を上げるのが目的ならば、意図と結果を対比するだけでは不十分で、それを最終目的である効率向上に結びつける方法を見つけないければならぬ、ということである。この議論はかなり重要だ。というのは、仮にこの順序を逆にして、効率向上に結びつける方法を見つけるのが評価の目的で、そのために当初の意図なるものを明らかにするのだ、とすれば、もしかしてその方法が別のやり方でみつけれられるのなら、こんな七面倒臭いことはしないでよい、ということになる。そんな便利な方法があるのならば、まずそれが先決で、それが見当たらない、考えつかない場合に、この方法による、よらざるをえない、というのが順序になる。先に財団のマネジメントが、一般企業のそれと選ぶところがない場面を指摘したが、それどころか実はこの問題提起、つまり、そんな便利な方法があるものだろうか、というのは、一般企業において早くから「品質管理」の問題として意識されていたものとはほぼ同一であるといえる。

意外に思われるかもしれないが、品質管理とは、「買い手の要求に合った品質の品物またはサービスを経済的に作り出すための手段の体系<sup>※2)</sup>」であるのならば、不思議でもなんでもない。「買い手の要求に合った」というのを、「求められる」と読み替えれば一目瞭然だろう。さて、第二次大戦後のわが国における品質管理の歴史は、大別して、無検査の時代・検査の時代・品質管理の時代・品質保証の時代の四時代に分類されるという。このうち、検査の時代というのは、良い製品を顧客に提供すべく、完成品を検査して分類し、一級品のみを市場に出そうというもの。それに続く品質管理の時代というのは、工程で品質を作り込んでいこう、つまり、出来上がった品物ではなく、製造工程を丹念に吟味していくことによって、より良い結果を期待するというスタンスである。ちなみにこのステージは昭和三十年後半にはもう終わって、次なる品質保証の時代に入ったといわれている。われわれがここで論じている個別プロジェクト、あるいはプログラムの評価というのは、いわば検査の時代に相当し、三・四十年歴史はさかのぼる訳だ。

財団のマネジメントが、一般企業のそれに比べて三・四十年遅れているなどと飛躍するつもりはないが、終わってしまったプロジェクトをつかまえてどうこうするよりも、生産過程、つまりプロジェクト・フォーメーションを精査することによってより良い成果が期待できるのなら、そのほうが望ましいのはいうまでもない。まして、この品質管理の時代でさえも、今日から見れば製造部門中心で、全部門がよりつどって品質向上に取り組む顧客本位の時代の前史である、と見られているのであれば、なおさら、この時代は早く卒業しなくてはなるまい。

※1) 片山正夫『助成財団のプログラム評価』公益法人1995年5月号14頁。

※2) 塹江清志・澤田善次郎編『生産管理総論』116頁 日刊工業新聞社（1995）。なお以下の品質管理に関する記述も同書による。

## 第10章

成果評価にあたって、計画と成果を対比させる必要があるのは、学習効果を期待する場合である。そのためには、個別のプロジェクト評価に代えて、プロジェクト・フォーメーションのシステムを確立することの方がより望ましい。これは一般企業の生産管理手法に学んだ結論だが、財団への品質管理手法の適用はいかにして可能だろうか。

### ・ 成果についての評価（その3）

製品製造の工程管理、つまりプロジェクト・フォーメーションについてはさまざまな機会に、さまざまな人々によって知見が披瀝されている。それらを網羅的に紹介するのが本稿の主旨ではないことから、ここでは代表的なもの二つの概要を述べて、おおよその感触と、これまでに見てきた「評価」作業との関係を明らかにするに留める。

ひとつは主として開発援助計画のために、国際開発高等教育機構（FASID）によって策定されたプロジェクト・サイクル・マネジメント（PCM）と呼ばれるものである。これはまず 参加者分析、つまりプロジェクトの対象とする人々（組織）の特性を把握し、次に 問題分析、すなわちこれからプロジェクトを考えようとする領域の中心的な問題を把握設定する。さらに、目的分析として、の問題が解決されたときの「望ましい状態」を示すことによって、それに到達する手段を明らかにする。そうすると、いくつかのプロジェクト案が作成されることになり、その中からプロジェクトの選択を行うことになる。ここまでの作業を行うと、(a)プロジェクト目標と同時に、そのプロジェクトが貢献することになる。(b)長期的な目標が明らかになる。当然、(c)成果が何であるかも定義され、同時にその成果を実現するための(d)活動内容も具体的に明らかになる。この四つの項目のそれぞれについて、目標値を設定し、それを測定する手段を明示する。さらに、外部条件とプロジェクト実施の前提条件、そして資源投入計画をたてる。これを一表にしたのがプロジェクト・デザイン・マトリックス（PDM）と呼ばれるものであるが、これが作ってあれば、効率性、目標達成度、もたらされた変化、当初目標の評価時における妥当性、効果の持続性、の五項目にわたる評価が自動的にできるとされる。<sup>注1)</sup>

PCMの味噌は、プロジェクト構成要素と前提条件の系統的な記述によって、評価対象とその手法を操作的に定義しているところにある。だから、成果評価にとどまらず、意図・方法論も評価対象に組み込まれているが、それらを相互に関連づけて一体として扱うことによって、極めて具体的な品質管理の方法を提示することに成功している。経済・医療・教育といった開発援助の主力分野、特にプロジェクト成果の定量的把握がしやすかったり、代替プロジェクトが明確に提案できるような場合には極めて有力な手段だろう。要するに、プロジェクトのねらいや効果が一義的で、スパッと割り切れる類のものには強い。ということは逆に、そのあたりが曖昧模糊、と

いってはなんだが、多義的解釈を許すような場合には、切れ味鋭く一刀両断、という訳にはいかない、ということになる。先の評価類型の中で、定性的（qualitative）評価といわれるものは、実はこのあたりの悩みを解決すべく工夫されたものだといってよい。例えば、冷戦構造終結後の世界秩序について、諸賢のご意見拝聴のシンポジウムを開きました、途上国十カ国の子供たちを集めてサマーキャンプをしました、の類をPDMで整理しようとしても、ちょっとしっくりこないように思います、という話である。ならばこの種プロジェクトの品質管理はお手上げかということ、そうしたものでもない。

ドラッカー財団が開発した“The Five Most Important Questions”<sup>注2)</sup>はなかなか良くできた評価マニュアルで、この種のプロジェクトを扱う財団の「製造工程」管理にも汎用性がある。五つの質問というのは、われわれの使命（仕事）は何か。われわれの顧客は誰か。顧客は何に価値をおいているか。われわれはどんな成果をあげたか。われわれの計画は何か。というもので、それぞれの質問は、三つから五つのワークシートに細分された質問を伴っている。例えば第一の質問は、以下の四つの質問に細分される。-1.われわれ（の組織）は何を達成しようとしているのか？ -2.われわれが求めている個々の成果とはどのようなものか？ -3.われわれの比較優位はどこにあるのか？逆に弱点は何か？ -4.われわれの使命は見直さなくてよいのか？といった具合である。五つの質問はこの順序に答えてゆく。つまり、この順序で系統だてて自分の組織と環境条件を整理していけば、最後の具体的行動計画（＝プロジェクト）にたどりつく、というあたりは先のPCMと同工異曲なのだが、一味違うのはドラッカーの独特の語り口が各ステップに反映されていることと同時に、この各ステップでは熟練した司会役（ファシリテーター）によるグループ・ディスカッションが予定されていることだ。そして、上の五つの質問を一見してわかるように、従来定量的な評価がしにくいと思われていたようなプロジェクトに対しても、極力それを試みてみようとしている、といってよい。

プロジェクト・フォーメーションについては、これ以外にも多くのマニュアルがあるが、要は、所与の条件だと考えがちないくつかの前提条件を、可変条件だという目で見直し、評価対象（意図・手段・成果）とその記述様式を何らかの枠組みで限定した上で、行おうとする（あるいは既に行った）作業をその枠組みに沿って分解して記述する。その分解して記述するプロセスそれ自体が、事業目的を再定義したり、見落としや矛盾点を発見するようにデザインされ、従って（特に不満足な）結果は何で、どの工程に原因があるかが一目瞭然になる、というしかけになっている。できあがった結果を、当初の意図や評価の物差しと対比してあだ、こうだ、というのではなく、いわば結果の「製造される過程」を分析して、それぞれのステップに無理がないか、改善の余地がないかを観察する。当然学習効果も期待できる、というものである。

注1) 『開発援助のためのプロジェクト・サイクル・マネジメント』（財）国際開発高等教育機構

注2) ピーター・F・ドラッカー編著、田中弥生訳『非営利組織の自己評価手法』ダイヤモンド社

## 第11章

成果評価が学習効果、つまり将来に向けてより良い事業を作っていくこと、をねらいとするのなら、個々の「終わった」事業を捉えてその結果をあげつらうより、その生みだされるプロセスに着目するほうが効率的である。それは単に効率的であるにとどまらず、不毛の議論を避けることにもなりそうだ。

### ・ 成果についての評価（その4）

何だか魔法の杖の一振りですぐ苦役から解放されたかのごとくであるが、もちろんそんなうまい話がある訳ではない。試してごらんになれば解るのだが、PCMであれドラッカーであれ、プロジェクトを計画するうえでイメージがはっきりしていないものは見るも無残なことになる。逆にいうと、「ちょっといいでしょう」と思いつきに走ったものが、本当に良いものならば、てだてを踏んで磨くことによって、きっちり仕上がるし、そうでないものは出直してこい。まさに「学習」効果あらたかな道具というのが真骨頂である。

ところが、という話に入る。品質管理を志す、いわば初学者財団にとってはここまでのアウトウで十分で、以下の議論はいらざる混乱を引き起こすものとなるから読み飛ばしていただいたほうがよい。また財団の業績は、どこかで検察官のような人が理非曲直を明らかにして「評価」せねばならぬ、とこれまでの議論に納得ゆかぬ向きも、以下はご遠慮いただいてもさしつかえない。

と、持ってまわった言い方をしているのには訳があって、所謂「分析的」経営手法には限界がある、どうかすると欠陥がある、というのが最近の経営学の通説になりかかっているからである。それも、猿のたまねぎの皮剥きではないが凡人に分析的手法を教えると淫してしまって、際限もなく細部にこだわって大局を見失う、という程度の話ではなく（ただし上記の手法、特にPCMについてその傾向なしとしない）、どうかすると財団の本質論にかかわりかねない話なのだ。

「分析マヒ症候群」（野中）<sup>注1)</sup>といわれるものはその一つで、これは要するにたまねぎの皮剥きの結果、まわりの求めているもの、つまりプログラムの効用に対する配慮や適応がお留守になる状態を指す。もちろん「意図」についての評価を伴っていれば避けられる事態であるが、几帳面に三段階の評価を行うことは必ずしも一般的ではない。評価作業をルーティンに落とし込んでいる組織さえ多くないことでもあるし、仮に意図評価を行うにしても、何年かに一回の財団全体のプログラム方向性の評価の時だけで、後は成果評価のみ、というのがあらかただろう。

であるとすれば、これが由々しい事態なのは改めて説明するまでもないだろう。評価といい、プログラム策定といい、いずれも財団の資源を費やす作業である。その資源配分に際してこうした危惧があることには認識しておかねばならぬ。

さて、元に戻って、学習効果はどうでもよい。結果が出たかどうかだけが問題なのだ、という向きには、全く別の手法がおすすめた、ということになる。意図はともかく結果だけを測定したい、という

のはどんな場合なのか。

投資効果を測定したい、というのが最も一般的だろう。これこれのオカネを使ってあるプロジェクトをしましたね、それで結果はどうだったんですか。惨憺たるものだったのか、まあまあ、それとも大成功ですか。

大成功の場合には使い途は多い。財団あるいは民間非営利団体の事業成功例としてパブリシティにも使えるだろうし、その事業がナントカ賞を受賞するかもしれない。柳の下の泥鰌を狙うのもよいだろう。

惨憺たるものだったらどうなるか。そもそも惨憺たるありさまというのは、どんな場合なのか。講演会を計画しました。当日聴衆はゼロでした、というのは解りやすい例である。入りはそこそこだったが講演の内容がひどいものだった、というのもこれに近い。もっともこの場合には内容がひどかったかどうか、意見が分かれることもある。それについてはすぐ後に触れる。それよりも、惨憺たるものだからどうだというだろう。

なぜそんなひどいことになったのか、原因を調べて他日に備えよう、というのはここで問題にしていない（学習効果を期待する評価には別の手法がお勧めであった。為念）とすると、担当者の業績評価でもする以外使い途はなさそうだ。どうやら、結果が出た、つまり「こんなにうまくいきました」という発表だけが評価目的であり、成功したケースだけに評価対象としての意味がある、ということになる。もちろん、そんなに惨めな結果ばかり生んでいるようだと、別のところから声の出る可能性はある。「あなたの財団は、租税の優遇措置を受けていますね。つまり本来ならば納めるべきだった税金を、世のため人のためにお使いになるということで免除しています。それがこんなひどい結果に終わるのならば、やはり税金の形で国に納めていただいて、優秀で有能な役人が使ったほうがよいようですね。」これは直接に本稿の主題ではないから深くは立ち入らないが、二つの点だけは指摘しておきたい。

一つは、失敗の申し開きができること（accountability）の大切さである。後ろ向きに「言い訳」を探して汲々とすることを勧めているのではない。失敗（失敗であるかどうかをどう判断するかについては後に述べる）がどのようにして発生したのかを開示し、それが組織内の約束ごととどのような関係にあるのかを明らかにすることの重要さである。オマエがやったから失敗したんで、オレだったらもっとうまくやる、とか、失敗しても別に咎め立てされるいわれはない、というのは別の話である。

というのも、第二に、失敗は必ず起こるものであり、百発百中成功オンリーというのはむしろどこかおかしいからである。成功の確度を上げることも評価の目的の一つでありうるが、決してそれだけが目的なのではない。失敗を目的とするプロジェクトというのは考えにくいだろうが、失敗の存在それ自体にも、他日に備えての学習効果以外には意味がない、と即断するのは早計である。いついかなる時でも正しい、という命題よりもそうでないもののほうが情報量が多い、<sup>注2)</sup> というのと同じことである。

注1) 野中郁次郎、竹内広高『知的創造企業』東洋経済新報社（1995）

注2) 小原誠『ボバー』p128 講談社（1997）

## 第12章

将来に向けての学習効果ではなく、単に何が起きたかだけが問題になる場面を見てきた。それはいきおい成功か、失敗だったかを述べるようになる場合が多そうだ。特に失敗したケースについては、いくつかの問題がありそうである。今回からはそろそろ「まとめ」に入る。

### ・ 成果についての評価（その5）

ここでは極端な議論をしている訳で、失敗から何も学ばない、というのはかなり非現実的な設定である。つまり、結果評価しか行わなかったとしても学習効果はもちろん期待できると考えて間違いはない。それどころか、多くの場合、当初の意図だ、結果との対比だ、という七面倒な作業より、結果として何が起きたか「だけ」を把握するほうがはるかに簡単だから、小規模な組織やプロジェクトの場合はむしろお勧めでさえある。ここで気をつけなければならないのは、「自画自賛」に陥らないように、つとめて具体的に結果を記述する態度である。万歳々々の大本営発表では、上に述べた成功事例としての使用価値さえなくなるというものだからである。

さて、成果把握にあたって実務的に悩ましいのは、成果把握が「一義的」にできるものではなく、価値判断を伴わざるを得ない場合である。例えばある問題について、何人かが集まって解決策を提言しました、というようなケースだ。提言書ができてはみたものの誰も見向きもしませんでした、実行もされませんでした。いったい、

提言書の内容がお粗末だったのか、内容は結構なのだが、  
-1. それをしかるべき筋に読んでもらえなかったのか、  
-2. 目を通してはもらったものの何らかの理由でとりあげられなかった、実行されなかったというのか。つまり「何が」うまくいかなかったのか、の特定作業である。成果評価は結果オーライのはずなのに、ここであえて七面倒な議論が始まったのは、「ここまでは」うまくいったのを確かめようとしている訳だ。

このうち については第3章の「意図の評価」で先見性の問題をめぐって、また第7章の「方法論の評価」のところで「馬鹿げた事業を発見するため」の評価として触れたのでここでは繰り返さない。ここでアウトならば、一巻の終り。これをクリアして に進むとどうなるか。

一番簡便なのは、 の部分は別の「事業」だと考えることだろう。そうすればこれまでに議論してきた考え方や手法がそのまま適用できる。観念的には分離できるが、実際には一体にならざるをえない場合はいささか厄介なことになる。例えば、ある事業が「それとしては」やたらにコストがかかり、とても納得をうるにいたらない。しかし、これはモデルで、初期投資は大きいものの、成功の暁には周辺に拡大していくことが期待でき、それならば十分に説得力がある、というような場合である。一般製造業の設備投資だったら冗談のような話だが、確信犯が多いこの世界では大真面目な議論になる。予期したようにはまわりに拡大していかなかったのが、「早く来すぎた」アイデアだったからなのか、それとも単に「おめでたい」夢物語に過ぎないのか、という話については、結果が出るまでの時間軸の問題として第4章を参照されたい。かくして「ここまでは」うまく

いったのだ、と成果評価は納得することになる。

### まとめ

財団における評価の試みにはいくつかのねらいがある。それを明確に浮き彫りにするために、プログラムやプロジェクトを「意図」、「方法論」そして「成果」の三ステージに区分し、それぞれの場面で、評価にはどんな意味があり、何をねらいとしているのかを明らかにした。

「意図」の評価のそもそもの目的は、「善き」意図であっても結果が仇となることのないようなアセスメントに始まる。しかしさらに一步を進めて、その意図があまりに先駆的なものであって、同時代の人々の支持を受けられない場合にどう対処するか、という点が問題になる。しかし、一般的かつ平均的なもの見方だけが大事なのだろうか？あえて「俗な」とはいわないまでも、手慣れた常識から自由で、のびやかに振る舞う余地、そのような自由度こそが民間非営利団体の存在理由であり、担うべき機能だとすれば、意図の評価にあたっては四角四面な態度を取るべきではあるまい。とはいえこれは諸刃の剣になりかねない。ひとりよがりの脳天気と先見性の区別はどうやってつけたらよいのか？ここで質の高いプロフェッショナルリズムの出番になる。意図の評価はスタッフの資質の評価に連なる、ということだ。

「方法論」の評価は「意図」についてのそれと比べると、ずっと実用度に対する配慮が強くなる。あの方法よりはこれがよい、という改善改革こそがねらいであって、それを忘れた「ないものねだり」つまり、最高至善を求めてのアラさがしには何のメリットもない。ということは、代案・対案を（たとえ明示の形でなくとも）示唆する評価であるのが望ましい。試験の答案を採点するだけの作業ではないということだ。

「成果」評価の要諦は、結果のもたらした効用を知りたいのか、あるいはよりよい結果をもたらしたかったのかを峻別するとことにある。前者の把握は効用と測定手段の定義さえはっきりさせれば比較的容易である。後者であるのなら、出来てしまった製品の質を論じるのではなく、製品製作のプロセスを管理して、よりよいものにするのが作業の目的になる。これは手段に拘泥して目的を忘れる危険性がないとはいえないものの有用な評価手法である。とはいえ実務的な見地からは、結果の善し悪しが意図によるのか、手段によるのか、甚だしい時には何が結果で何が手段なのか解らなくなることがないとはいえない。でもゆっくり考えてみれば、もつれた糸はほぐれるものだ。

延々と説き去り説き来たった結論は、なんのことはないわずかそれだけのことだったのか、といえども蓋もない話だが、英語の論文風にこれまでの Executive Summary をつくればこんなことになるのだろうか。しかしこれには深い（と筆者自身が言うのは面映ゆいものがあるが）含意がある、という点を再度強調しておいたほうがよいだろう。

## 第13章

「どのように」「何を」評価するか、については一通りの議論が終わった。

ここでは「なんのために」評価をするのか、という原点に立ち戻る。

いうまでもなく、本稿は財団の（といているが、実は民間非営利組織一般を意識した議論をしているつもりである）評価について述べている。

つまり、その製品やサービスが市場によって評価を受ける訳ではなく、ひとりよがりになったり、非効率的な経営に陥る可能性を内在させている組織において、評価がどのような意味を持つか、評価にいかなる位置づけを与えるのが適当か、ということである。

この設問から既に明らかのように、ここで問われているのは、営利企業が市場で受ける評価と同じ機能を財団が自ら行う評価に期待できるか、という点が一つ。答えがノーなら“Then what?”ということになる。

さて、答えは明らかにノーである。事実上強制力のある市場評価は、少なくとも理論的には資源の最適配分に役立つ。それに比していわば恣意的で任意性の強い自己評価（外部機関に依頼しても所詮は自己評価である）はまるで違う。全知全能とはいわないまでも公平無私、判断力に優れた経営者が行なったとしても本質的には変わらない（第一そんな経営者ならもともと評価など必要としないだろう）

同じく市場原理に馴染まないサービスを提供する公的セクターは、成果評価から免責されている。それは国民から（法律という）明示の形で負託を受けているからであり、その担保として、あまりに低い顧客満足しか与えられない公的セクターは、例えば次回選挙において支持を失う。それに対して一種黙示の支持しか受けていない財団は、何を拠り所にしてサービスの質を保證するのだろうか、そのために自己評価がどう機能するのだろうか、と考えてもよい。

ここで議論に内在する二つの要素が明らかである。一つは判断基準として参加者に「受け入れられた」ものがあるかどうか。第二には提供するサービスの質を吟味する手段の有無である。

市場原理が質を保證するとは限らない。最近のテレビの俗悪番組を見ても、かつてのVHS対ベータマックスを見てもそれは明らかである。しかし、一方で市場参加者（需要・供給双方とも）はそういうものとして納得づくで参加しているし、他方ではそれゆえにこそ市場原理を万能としないサービス提供が求められる。

納得づくという点では公的セクターの法律による負託も同じことである。そうすると財団（あるいは民間非営利組織）の提供するサービスのみが評価についてのインフラを欠いているのだろうか。あるいは黙示のものである結果、さまざまのものが混在して、一義的に定まらないのだろうか。

この問いかけの中に既に回答が用意されている。

つまり、納得づくであれば質を問われることはない。もちろん俗

悪番組やお役所のむだ遣いなどが批判されることはある。しかしそれは個別のいわば異常値に対するものであって、仕組みそのものに向けられることは少ない。

言葉を換えれば、仕組みが認知されているところでは、品質管理は（組織の）内部問題であって、それ以上のものではない。ちなみに、お役所仕事の民営化必要性が望ましいとされるのも、市場という仕組みの品質管理の方が、お役所内部のそれよりも優れている、という経験的知識が存在するからである。

とすれば、財団の存在、あるいはその社会的機能そのものが、どのような形で世の中に納得されているかをまず問うてみる。それに肯定的な回答が用意されていればそれまで。否定的ならば、製品の質についての検証手段が問われることになる。

財団の社会的認知度についてはこの小稿のよくするところではないから詳論はしないが、とりあえず否定的に解しておこう。

そうすると、ここで初めて財団の提供するサービスについての質が問われることになる。無罪（質が良い）が証明できないかぎり有罪（ロクなものではない）というのはあんまりだから、その逆であるとしてしよう。

個々の製品についてそのチェックをするのはあまりにも膨大な作業であるし、非効率である。だから、ロクでもないものを生み出さない仕掛けについての検証とその公表が財団の評価作業の要諦だと考えれば、自ずからその方向性が明らかになる。

のみならず、評価作業に向けられるエネルギーを、より積極的で意味のあるものにするためには、「納得」、つまり社会的認知度を上げることこそがより重要になる、ということも浮き彫りにされるはずである。

これまでに財団の事業評価について述べてきたことは、すべてこの枠組みのなかでなされたものである。

その視点から振り返れば、お役所の指導監督が財団の仕事の質を向上させるという議論がいかに的外れかも明らかだし、善意と志さえあればすべて許されるとするのも適当ではないことも同様に明らかになる。

アマチュアが休日を楽しむ草野球とプロ野球を比べてみればすぐ解るように、専門性、目的、さらには影響を及ぼす範囲によって、要求される品質（あるいは資質）は全くといってよいほど異なっている。同時に、アマのゴルファーが目指すものと、プロのそれが違うのも明らかであり、それによって達成度についての評価も異なってくるのは当然である。まして「同じようにボールを打つただから」ゴルフと野球の評価は同じようなものだ、という議論もナンセンス以外のなにものでもない。

いきなり飛躍した結論になって恐縮だが、要はより良く社会的機能を果たすための道具の一つを開発しよう、というのが財団の事業評価の目的とするところであり、この小論が多少の役に立つことがあれば望外の幸せである。