

平成13年度経済産業省委託調査

技術評価に係る評価人材の育成等に関する調査

報告書

平成14年3月

財団法人 政策科学研究所

はしがき

本報告書は、平成13年度の経済産業省の公募調査事業『技術評価に係る評価人材の育成等に関する調査』を財団法人 政策科学研究所が実施した成果をとりまとめたものである。

我が国では、近年、評価の時代を迎え、研究開発を含む対象として様々な評価が行われ始め、プロジェクト評価が広く浸透する一方で、プログラム、制度、政策のような、より複合的な評価対象にも検討の視点が広がりつつある。これに伴い、評価作業にかかる負荷は増大し、また、評価パフォーマンスもより高度なものが求められている。

したがって、評価の信頼性を維持・向上させる上で、これを担う評価人材を量的質的に確保するという課題が急速に顕在化しつつある。事実、平成13年3月30日に閣議決定された第2期科学技術基本計画（以下、第2期基本計画）において、「優れた成果の創出・活用のための科学技術システム改革」の一つとして、「評価システムの改革」があげられている。この中で、競争的な研究開発環境の実現と効果的・効率的な資源配分に向けて、「評価に必要な資源の確保と評価体制の整備に重点を置いて改革を進める」としており、具体的には「研究費の一部を評価の業務に充てる」などして評価体制の充実を図るとともに、「研修等を通じて人材の養成に努める」ことが挙げられている。

本調査は、評価に関連する人材を効果的に育成するために、評価人材の類型別に中長期的なビジョンと施策の提起を行うことを目的とする。このため、研究開発評価に対して先進的な取り組みを行っている欧米における育成方策の動向調査を実施したうえで、我が国にふさわしい評価体制（評価運営及び研究評価者）を構築するための基礎的資料を得ることを試みた。本調査が、我が国の評価人材の適切な育成システムの構築にいささかでも貢献することがあれば深甚である。

末尾ながら調査にご協力頂いた専門委員会ならびに有識者の方々に心から謝意を表したい。

平成14年3月

財団法人 政策科学研究所

目 次

はしがき

目次

要旨

調査の目的・内容・方法・体制

序	1
第1章 評価人材の養成と受用の構造	2
1. 1 人材受用側の全体像	3
1. 2 人材養成側の全体像	5
1. 3 必要とされる評価の教育・研修システムの全体像	7
第2章 何を修得すべきか	10
第3章 カリキュラム内容の概要	12
第4章 対象者の区分	16
4. 1 評価マネジメント人材	16
4. 2 社会経済性分析エキスパート	16
4. 3 社会経済性評価レビューア	17
4. 4 科学技術ピアレビューア	19
第5章 研修コース案	23
5. 1 マネジメント・コア・コース（本格的入門コース）	24
5. 2 エキスパート・コース（発展コース）	25
第6章 転換コース案	39
6. 1 マネジメント支援者養成コース	39
6. 2 エキスパート養成コース	45
第7章 高等教育専門課程	46
7. 1 科学技術経営政策研究課程（大学院修士課程）	46
7. 2 大学付設の期研修コース	46
第8章 評価人材整備のための施策	49

<参考資料>

参考資料 1. 事例	55
事例 A 評価人材養成プログラムに関する事例	55
A-1 マンチェスター大学 P R E S T における教育コース	55
A-2 オランダ・トゥエンテ大学における研究評価の教育コース	62
A-3 米国ジョージア工科大学 (G I T) における教育コース	66
A-4 A A A S 科学技術フェローの オリエンテーションプログラム	78
事例 B 評価人材の確保・活用に関する事例	85
B-1 米国 N I H におけるレビュー体制と 評価人材に関する議論	85
B-2 欧州の行政機関等の評価とその人材育成の取り組み	100
B-3 E U における科学技術ピアの公募・選定・行動規則等	110
参考資料 2 欧米の評価研究者・エキスパートとネットワーク	123
2. 1 評価研究者・エキスパートの国際的ネットワーク	123
2. 2 米国の行政内部の研修コース講師候補	135
参考資料 3. 評価人材が学習すべき重要文献	137
3. 1 評価に関する主要総説文献	137
3. 2 ピアレビューに関する重要総説文献の事例	138
参考資料 4. 現在の評価関連人材に関する問題点の指摘例	167

要 旨

研究開発評価が推進されるに伴い、評価の質を担保できる評価専門家の欠如が各所で認識され始めた。評価の専門性の養成と普及を放置するならば、評価に対する信頼性が醸成されないまま形式的な評価システムが運用され続けることになりかねない。また、評価は戦略策定とともに、我が国の行政運営の新しい局面を拓くうえで不可欠な能力となってきた。

しかし、評価人材の専門性を高めることは、その種類により異なるものの、時間を要する課題である。評価人材の養成を戦略的に始める必要がある。評価先進国でも評価人材が相応の厚みをもつには持続的な強化の取り組みを必要としたのである。我が国には評価に関する能力やスキルが組織に定着していないので OJT には制約があり、また高等教育機関を整備して相当する評価人材が生み出されるのを待つには時間的余裕がない。20 年来の課題に追いつくための短期の速修体制を我が国に適合する形で編成整備し強力で推進すること以外に道は無い。本報告書はその内容を具体的に提言するものである。

第 1 章 評価人材の養成と受用の構造

1. 人材受用側の全体像

評価に関わる人材に期待される専門的機能は、1) 評価システムを設計し運用するマネジメント機能、2) 評価過程において評価自体を担う機能、3) 評価の前段として評価対象や状況等を分析し全体像を明らかにする評価支援機能、の主に 3 局面にわたる。

評価先進国では、評価研究者は大学等の高等教育研究機関に、実務的専門家は、行政やファンディング機関ないし研究機関等の評価実施部門、さらにはシンクタンク等の評価支援機関、民間企業や国立研究機関等の研究機関の内部にも集積されているが、我が国では極めて薄く分散配置された状況にある。

2. 人材養成側の全体像

本来、評価人材は、大学等の高等教育機関から育成・供給され、上述のような評価人材の受用機関の内部に集積されているべきものであり、欧米先進国ではそうになっている。重要なことは、評価実施機関の内部で OJT のメカニズムにより評価の実務的専門家の拡大再生産が行われていることであり、研究開発の本格的な経験者が不足して

いる行政機関等に対しても高等教育機関以外にも専用の研究者転換コースを用意している国もある。

しかし、我が国では明示的な評価システムの歴史が浅く評価専門人材のニーズが少なかったこともあり、高等教育機関には研修コースや転換コースも含めて該当専門コースが設置されず、評価人材のストックは殆ど無く、少数の研究者や実務的専門家が散在しているのみである。

3. 必要とされる評価の教育・研修システムの全体像

我が国の現状を踏まえ、適合する教育・研修システムを構想する必要がある。

1) 獲得すべき能力内容（コンテンツ）

我が国の高等教育カリキュラムには評価の背景的な素養を扱う講義科目もほとんどなかったことに留意し、教育研修の内容としては、評価法など評価のみに局在化させるべきではなく、その背景となる知識やスキルを含めておく必要がある。

2) 養成すべき評価者とその機能（アクター）

評価関係者としては、①評価システムを運用し、その信頼性やパフォーマンスに関わる、評価のマネジメント人材、②評価のより困難な領域を担い、評価を支援する人材、例えば、社会経済性分析エキスパート、③社会経済的側面に関わる評価を行う社会経済性評価レビューア、④科学技術の質に関わる評価を行う、科学技術ピアレビューア、に大別できる。とくに、①②③の評価人材の養成・確保、能力開発が我が国にとっての課題といえる。

3) 教育研修システム（システム）

我が国で今後整備すべき人材養成システムとしては、①評価マネジメント人材や調査分析エキスパートを擁しているべき機関の評価人材の能力向上を図る短期研修コース、②研究開発者がその研究開発経験を活かして評価人材に転換するための短期転換コース、③高等教育機関で行う本格的な専門家養成コース、の3種類が考えられる。

以下、この研修内容について第2章～第7章で具体的に展開した。

第2章 何を修得すべきか

研究開発評価の実務では、一定の大きな枠組みを理解し踏まえる必要はあるが、具体的な内容までをマニュアル化することは殆ど不可能であり、また意味がない。研究評価は全て「応用問題」であり、それを目的とする場の環境と文脈に合わせて「解決」する能力が必要である。このため修得すべき能力として、次のような構成が基本とな

ろう。

① 評価のための実務的スキル

ワンセットの実務的方法論

② 手法を使いこなすための基盤的能力

実務的方法論を使いこなすための基盤的能力

③ マネジメント全体を見通す広い視野

評価をマネジメント全体の中に位置づけるための能力

第3章 カリキュラム内容の概要

カリキュラムによる修得目標としては、第一に、研究開発評価の推進体制の全体像が把握され、第二に、評価の諸局面についての理解が深められることが必要である。その上で第三に、評価の方法論が把握されなくてはならない。研究開発評価に特化した研修は、しばしば評価の捉え方を硬直的にし、あまり生産的ではない。評価論全般の知識を踏まえた上で、研究開発評価の特殊性を考慮した方法論について理解を深めることが必要である。さらに第四に、評価推進体制を策定し設計し、また、運用するための導入的な枠組みの理解が必要である。その際、評価システムとしての重要な局面やポイントが押さえられる必要がある。

第4章 対象者の区分

評価システムを担う人材の重要な4類型と人材養成上の対応策の特徴を指摘した。

1. 評価マネジメント人材

大きく2種類あり、1) 評価システムを統合的に管理運営するための評価支援部署に配置されている、評価マネジメントのコア人材と、2) 評価実施部署に配置されて担当課題の評価の運営に当たる、研究開発評価マネジャーがいる。後者に対して入門的な評価マネジメント研修コースを提供するとともに、前者には加えてさらに発展的な研修コースを提供する体制をとり、評価マネジメント人材の急速な充実を図ることが望ましい。

2. 社会経済性分析エキスパート

研究開発の評価において経済的ないし社会的な側面の認識は、科学技術的側面に比較し極めて困難である。そのために統計資料や様々な調査結果を専門的に分析し、状況認識を深めることを支援するエキスパートが必要である。課題と状況に合わせたオリジナルな分析が必要であり、これを担える実務的専門家を養成する必要がある。我

が国では現在その任に堪える人材は極めて限定的にしかない。入門的な研修コースに加え、発展的な研修コースを提供することが必要である。とくに海外で深められている事例によるケーススタディ、新しい観点やアプローチを付加する先進的研究成果の学習が効果的であろう。

3. 社会経済性評価レビュー

社会経済的な評価は、その側面を扱う独立の評価パネルか、科学技術的側面からの評価パネルと統合したパネルで実施される。評価パネルは、社会経済性分析のデータが存在する場合とそうでない場合とが想定される。

分析データが整備された状況下での社会経済性評価レビューには、シンクタンクや行政に蓄積される当該関連分野の経験あるシニア・エキスパートを登用すべきだが、その存在は現在極めて限定されている。そのため、これまでは科学技術人材の中から経験豊富なセニオリティや著名な人材であるエminentによるパネルを構成して委嘱してきたが、実質的に深い評価は困難であった。社会経済性分析エキスパートの育成が望まれている。

分析データを前提としない比較的単純な枠組みでなされる社会経済性評価レビューは、企業内の事業部門での企画や新規事業部門の専門家（研究所内だけで活動してきた人材はしばしばシーズ側からの評価に偏している）、ベンチャー・キャピタリストなど企業立ち上げの専門家や地域での企業コーディネートの専門家などに、広くしかし薄く存在しており、今後ともその直観的能力に判断を依存すると共にこの強化も必要である。

4. 科学技術ピアレビュー

科学技術ピアレビューは、科学技術の質に関わる評価を行うために、当該専門分野に通じた研究者・技術者によりなされる。科学技術ピアレビューは、専門的な深さと同時に高い見識や広い視野を備えた人材が適しており、そのリストは、評価の実績を通して、また推薦や公募を含めて、更新充実させるべきである。なお、我が国では少数の評価者への負担がまして「評価疲れ」現象のあることや、学術の細分化の中で利害関係を持たない評価者が希少化する問題、研究者に評価活動に対するインセンティブや責任ある関与が弱まってきた問題もあるが、研究開発コミュニティ基盤の健全性にかかわるものとして実態の把握や対応策を検討し始めるべきである。

評価マネジメントにおいては、科学技術ピアレビューを選抜し、動機づけ、評価目的に沿って、その役割を限定して評価能力を十全に発揮してもらうことが肝要である（一方で、他の評価人材へのキャリア転換を支援することも必要である）。その際

に、評価課程の情報の開示や「評価者の評価」などを実施して適切な緊張感のある役割を果たす環境を整えることが有効である。ピアレビューの研修コースは特に設定する必要はないが、評価マネジメント側は適正なガイダンスを行うべきである。すなわち、評価システムの基本的な枠組みや評価者の位置付け、期待される能力について、また、評価のクライテリアについて、十分な理解を促すことが必要である。さらに、評価制度の規程や評価者の倫理的養成等に関して具体的に注意すべきである。

第5章 研修コース案

評価関連部署に在席する職員を対象とした研修コース案として、本格的な入門コースである「マネジメント・コア・コース」と発展的なコースである「エキスパート・コース」を検討した。

1. マネジメント・コア・コース

本格的な入門的コースであり、評価対象に相応しい評価システムを設計し運用できる評価マネジャーの養成をめざすものである。行政機関内部の評価実施担当者や研究機関等の機関評価のマネジメント担当者等がこれに該当する。またこのコースは同時に、コア・マネジャーやエキスパートのための基盤的な入門コースでもある。

第3章で検討したカリキュラム内容のほぼ全域に亘って行われるべきである。その目的は、技術経営マネジメントに関わる知識の修得（日本の研究開発システムの理解を含む）、研究開発評価システム全般に関する知識の修得（日本の評価システムの実態及び欠陥や修正方向の理解を含む）、そして研究開発評価のマネジメントとメトリックス及びそれらを統合する方法論についての全体的実務的な運用スキルの修得である。

我が国の評価の特殊な事情を深く理解し踏まえた上で基本的な評価制度が設定されるべきで、我が国の状況に通じた国内講師による研修が望ましい。研修期間は5日程度が妥当である。

2. エキスパート・コース

一定の評価実務の経験を積んだ人材に対して、さらに発展コースとして用意するものであり、評価のための調査分析を担える「評価エキスパート」の養成をめざす。我が国には教材として適切な事例が集積されていないので、「海外の実践事例に学ぶシリーズ」と「海外の先進的研究に学ぶシリーズ」を設定し、海外講師を招聘して行うことが有効であろう。なお、研究に学ぶことに関連して留意すべきことは、研究論文の殆どは研究者の提案にとどまるもので、限定された課題の理解を深めることに焦点

をあてて活用すべきことである。

1) 海外の実践事例に学ぶシリーズ

我が国で必要性の高い分野に焦点を置くべきである。したがって第一に、社会経済性分析に関わる課題を扱う。独立大型プロジェクトの事前評価とプログラムの途上・追跡評価が2大課題である。これらを扱う専門家は欧米でも行政内部に少なく、得意分野を持つ外部シンクタンクに適宜依拠している。研修内容は、その事例を主題としたケーススタディを通じて、広範囲な分析やアプローチの多様な構成・展開での専門性の深さの紹介を受ける。Battelle 研究所の DOE の大型プロジェクト評価や SRI International の NSF のエンジニアリング関連プログラム評価が典型例である。1シリーズで5日間程度を要する。

第二に、R&D 評価システムを状況に合わせて設計・改善できるように、多くの事例や問題点を扱うもので、ベンチマークを行う。欧米でも行政内部でできない高レベルの設計や運営は外部の専門家を組織して対応することが多い。その重要事例を主題としたケーススタディを通じて、一般論でなく深い多様な経験に裏づけられ喚起される知識に触れることが重要である。AAAS による米国州政府の、テクノポリス社による欧州各国政府の、InDyne 社による NASA の科学振興プログラムの、それぞれの評価システムの設計と運営などが典型例である。1シリーズとして3日間程度を要する。

第三に、研究開発プログラムの下で展開されるプロジェクトの評価は、研究開発関連省庁の評価の大半を構成している。研究開発プログラムの運営に関して先進事例のベンチマークを行う。多様なプログラムの運営や評価ポリシーの違いを比較検討することを通じて、今後のプログラム運営の高度化に有益な情報を生み出すことが期待できる。NSF、NIH、DARPA、NASA、NOA などの米国リサーチ・エージェンシーにおける研究開発プログラムの運営や、DTI、EPSRC などの英国の研究開発プログラムの運用が典型例である。経験豊かな実務者を講師に1シリーズ3日間程度を要する。

なお、機関評価は機関ごとに極めて個性的であるはずであり、機関評価の事例紹介はあまり意味がない。むしろ本格的入門コースの中で機関評価のために必要な基盤的知識と共に、機関の個性を相対化した知識として取り扱われることがふさわしい。

2) 海外の先進的研究に学ぶシリーズ

欧米では、科学技術政策や技術経営などの広い専門性をカバーする研究者群に加えて、研究開発評価という、より専門的な分野に特化した研究者が集積している。評価の中心的課題が社会経済性とこれを含めたマネジメントにあることから、多様なキャリアを持っている。近年、評価研究の蓄積をもとに専門性を強めた研究者たちは国際

的なネットワークを形成し相互研鑽を強めている。これらの活動の成果物自体が評価人材の重要な学習文献である。

このような海外研究者を講師とする研修コースを設定する上で注意すべきことは、一般論は序論的な時間に限定することである。彼等に期待するのは、評価に関する新しい整理された研究結果としての概念や知識を教育することである。実務的経験を相対化し論理化して普遍的知にしたものや、逆に、コンセプトを深めた本質的な知見は、実務的活動の上でも先導的な機能を期待できる。

第6章 転換コース案

転換コースは、研究開発人材等が、評価の専門性を身につけてキャリア転換し、評価関連部門に配置されることを前提とした研修コースである。我が国では評価の実務的専門家が圧倒的に不足しているため、このようなメカニズムを強化する必要がある。

米国では既に 60 年代末以降に転換メカニズムが計画的に強化され、そこでの養成人材が科学技術行政の中核で活躍している。端的に米国 AAAS が主催する転換コースは科学技術行政に人材を送りこむように焦点が絞られている。欧州でも高等教育機関での柔軟な専門移動を背景にキャリア転換が図られている。

我が国に適合的な転換コースを2種類構想した。

1. マネジメント支援者養成コース

マネジメント支援者養成コースは、科学技術の研究者としての専門性を備え行政に対して様々な形態で支援する人材を想定し、配属前に高密度な集中研修を用意するものである。1) 現状の出向制度を自覚とインセンティブをもつ派遣者論理に立つものに置き換えるため、出向元で公募・選抜を行う「出向コース」と、2) 任期付ポストで行政内に中途採用する際に行う「転職・転換コース」がある。我が国では、期間限定的な支援者が専門性を活かして継続する、ないし官民の広い活動の場を得る流動的なメカニズムを整備する必要がある。AAAS 研修とその修了者の多様なキャリア・パターンは一つのモデルとなる。

カリキュラムでは、行政組織の概要、政策形成過程、予算策定・執行プロセス、行政の主要課題、派遣先の特殊問題などが扱われ、「転換コース」ではさらにキャリアアップ・メカニズムの紹介等がなされる必要があり、両コースとも2週間程度を要しよう。

2. エキスパート養成コース

エキスパート養成コースは、社会経済性分析のような専門家を養成し、その学問分野の専門性を行政内に導入することを想定して行うものである。分野の特性上、研究開発人材の転換研修コースと位置づけられ、上記のような「出向」と「転職・転換」の2コースを整備すべきである。このエキスパート養成コースの修了者から、さらに上記マネジメント支援者養成コースなどを経て行政内部で活躍するメカニズムも考えられる。このような転換コースは2～3週間の研修を要するが、年間数回、全省庁的にまとめて行うことになろう。

第7章 高等教育専門過程

科学技術の経営や政策を主題とする実務的な高等教育課程は大学院レベルであるが、欧米では既に30年ほどの歴史があり多様なコースが提供されておりその卒業生により評価体制に厚みをもって専門性の維持が図られている。我が国では、このようなコースは端緒的分散的に留まっており、本格的に整備される必要がある。しかし、人材問題への対応としては時間的にこのメカニズム以外の対策を急ぐ必要がある。

大学が提供するものに短期研修コースがある。評価人材育成のために、単位等の資格取得と無関係に2～3週間程度設定する集中研修コース、および単位・資格取得を可能とする形で学期に合わせて運用し休日ないし夜間に受講できるパートタイム研修コースなどが考えられる。

また、補助的にはインターネットを利用したeラーニングの活用等も考えられる。

第8章 評価人材整備のための施策

評価人材育成に関して、次のような内容やプライオリティをもつ施策を提起する。

1. 多様な研究開発評価人材に見合った養成メカニズムが必要

研究開発評価の多様な局面を担うために、専門化した多様な評価人材が必要であり、それぞれの評価人材に見合った多様な研修コースを目的に合わせて設定する必要がある。(表1)

2. 先ず「評価マネジメント人材」と「社会経済性評価人材」の養成から着手すべき

評価の状況と人材充足状況、職務の重要性等を考えると、コア人材のみならず評価マネジメント人材、全く不足している社会経済性評価人材、そして行政機関や資源配分機関の在籍者の研修に留まらず研究開発者からそれら人材へのキャリア転換の順に重点的に取り組みを始めるべきである。

表1 研究開発評価人材とスキルアップのために必要な研修コース

主な所属機関	評価人材の種類	マネジメント・コア・コース (本格的入門コース)	エキスパート・コース (発展コース)	
			海外実践 事例に学ぶ	海外研究 事例に学ぶ
行政機関・資源配分機関等	評価担当部署 (評価コア人材)	○	○	○
	評価実施部署 (評価一般人材)	○		
外部支援機関 (シンクタンク等)	エキスパート	○	○	○
研究機関	評価担当部署 (評価コア人材)	○	○	
	評価コアへの 転換人材	○	○	○
	行政機関等への 転換人材*	○	○	○
大学等	評価研究者		○	○

*この他に行政入門研修が必要

3. 研究開発評価に限定しない広い能力開発であるべき

評価の実務は、評価の目的や状況に合わせ、評価の在り方について創造的に取り組むべき「応用問題」である。修得すべき能力とそのための研修内容として、1) 評価のための実務的スキル、2) 手法を使いこなすための基盤的能力、3) マネジメント全体を見通す広い視野と知識、が必要である。

4. 研修内容に相応しい仕組みと講師の選定を

我が国の評価状況の課題と文化依存性、評価事例や講師適任者の蓄積を鑑みると、研修内容と講師の選定には特段の配慮が必要である。すなわち、入門的コースは我が国独自に対応し、また、高度な発展コースは海外講師陣を活用する基本態勢が妥当であろう。(表2)

表2 主な研修内容と妥当な講師像

研修コースの区分	主な研修内容	妥当な講師像
本格的な入門コース	<ul style="list-style-type: none"> ・評価の推進体制全般に関する理解 ・評価の諸局面の把握 ・評価の方法論の修得 ・評価推進体制の設計と運用 ・ケース・スタディと演習 	国内の評価事情に詳しい評価研究者や実務者
発展コース 海外実践事例に学ぶコース 海外研究事例に学ぶコース	<ul style="list-style-type: none"> ・独立大型プロジェクトの事前評価事例 ・プロジェクトの途上、追跡評価事例 ・研究開発評価システムの設計と運用事例の比較 ・従属型研究開発プロジェクトの事前評価事例の比較 ・最新の研究成果 ・実務的評価事例の相対化 	海外の評価実践経験の豊富な研究開発評価エキスパートおよび評価実践経験の豊富な研究者 海外の第一線で活躍している研究開発評価研究者
転換コース	以上の研究開発評価の研修内容の他に、行政メカニズムの理解のための研修内容を付加する	行政メカニズムの理解のためには、シニア行政官と政策研究者
大学付設のコース 短期集中研修コース パートタイム研修コース	基本的には、上記の「本格的入門コース」と「発展コース」の内容と同じ 上記「短期集中研修コース」の他に、その内容を使いこなし発展させるための「基盤的能力」の開発とマネジメント全体を見通す広い視野を獲得するための「経営学」関連学問	基本的には上記の「本格的入門コース」と「発展コース」の講師と同じ 上記の「短期集中研修コース」の講師の他に、システム論や経営学の専門家

調査目的・内容・方法・体制

1. 調査目的

我が国では、研究開発を対象として様々な評価が行われ始め、プロジェクト評価が広く浸透する一方で、プログラム、制度、政策のような、より複合的な評価対象にも検討の視点が広がりつつある。これに伴い、評価作業にかかる負荷は増大し、また、評価パフォーマンスもより高度なものが求められている。したがって、評価の信頼性を維持・向上させる上で、これを担う評価人材を量的質的に確保するという課題が急速に顕在化しつつある。本調査は、評価に関連する人材を効果的に育成等するために、中長期的なビジョンと施策の提起を行うことを目的とする。

2. 調査内容

評価を担う人材を類型化（例：科学技術の側の評価人材、経済社会／ニーズの側の評価人材、評価のマネジメント／支援人材）し、今後の評価システムの動向もかんがみて、その育成と活用に関する課題の整理を行う。これを踏まえて、関連事例を参照しつつ、我が国の実態に適合した、評価人材の育成のためのビジョンとプログラム等の提起を行う。

本調査は、文献資料の収集分析、国内関連事例調査、有識者・評価関係者ヒアリング調査、専門家委員会による討議、を通じて実施する。

調査項目は、次のとおりである。

(1) 我が国の研究開発施策等の評価人材における課題の整理

我が国の評価システム及び評価人材の需給やニーズの状況と動向を踏まえ、その育成と活用に関する問題点と課題を抽出・整理する。

(2) 評価人材の育成システム等の在り方

我が国で本格的な評価人材を育成するシステム等の在り方を検討し、関連事例等を参照しつつ、評価者育成のビジョンとプログラム等を提起する。

育成プログラムとして、例えば、人材類型別に、本格的な評価教育・研修プログラム（大学院レベルでの教育）、評価実務（行政業務受託機関）を通じた能力開発プログラム、研究開発人材の評価人材への転換プログラム等を検討する。

3. 調査研究体制

調査研究の専門性などに留意し、当研究所内に下記の構成の専門委員会を設置して指導と助言を得た。

<専門委員会>

委員長	平澤 冷	政策研究大学院大学教授
委員	丹羽 富士雄	政策研究大学院大学教授
	村上 路一	(株)宇宙情報技術研究所 副社長
	桑原 裕	(株)ホエールズウェブ・ドット・コム取締役
	西村 吉雄	東京大学大学院工学系研究科 教授 (平成14年2月15日まで日経BP社編集委員)
外部委員	林 隆之	大学評価・学位授与機構
	下斗米一郎	(財)松下政経塾

本調査研究を実施した当研究所内の体制は以下のとおりである。

<実施体制>

総括責任者	大熊 和彦	(財)政策科学研究所主席研究員・研究部長
	伊東 慶四郎	(財)政策科学研究所主席研究員
	猪瀬 秀博	(財)政策科学研究所主席研究員
	元川 浩司	(財)政策科学研究所主任研究員
	川島 啓	(財)政策科学研究所研究員
	斉藤 文子	(財)政策科学研究所研究員
	大谷 卓史	(財)政策科学研究所客員研究員
	上田 昌文	(財)政策科学研究所客員研究員

序

我が国の行政改革の課題は、組織改革から組織経営へその重点が移行しつつあり、同時に新たな組織経営を担うべき人材の養成が喫緊の課題となっている。組織経営の課題として先ず取り上げられたのが「評価」である。

研究開発評価については、他の行政改革に先んじて 97 年から本格的な検討が開始されたが、機関評価やプロジェクトの評価が進められるに従い、評価の質を担保出来る評価専門家の欠如が各所で認識されることとなった。

研究開発の評価について多少歴史的にその経緯を眺めてみよう。86年に始まる土光臨調の枠組みの中で、一度この課題が取り上げられ「研究開発の評価指針」としてその内容が閣議決定されたことがある。しかし、その内容は、研究機関に定着しないまま今回の改革のうねりを迎えることとなった。その原因は、一方では推奨された評価方法が当時の技術経営の状況と不適合であったためと思われるが、他方で多分に民間における研究開発費バブルの影響もあったと考えられる。本来ならば、オイルショック期を境として、キャッチアップ、高度成長の枠組みから、フロントグループ、低成長の枠組みへ行政のあり方が適切に切り替えられているべきであった。土光臨調の必要性はこのような背景を踏まえたものである。低成長であるならば重点化のための選択が必須であり、またフロントグループの一員であるならば、固有の課題を認識しそれに先見的に取り組むための体制整備が重要となる。必要な能力は「評価」であり「戦略策定」である。産業技術政策の有効性が後退していく歴史的経緯と、行政におけるこれら能力の獲得の遅れとは決して無関係ではない。

このように見てくると、今回の行政改革の中で唱えられたニューパブリックマネジメント、つまり「行政管理から行政経営へ」という流れの源流は、意外に古くからあり、かつて積み残した課題に改めて向き合いそれと取り組むものであるといえる。

組織に定着していない能力やスキルに関しては、もとよりそれを OJT で獲得することは不可能である。また 20 年来の課題に追いつくためには、短期の速修体制を強力に展開する以外に方法は無い。本報告書は具体的にその内容を提案するものである。

第1章 評価人材の養成と受用の構造

研究開発評価の大綱的指針の見直しや独立行政法人の評価委員会の発足、また省庁における政策評価の実施等に伴って、我が国での評価業務は質と量ともに拡充しつつあり、評価に携わる人材の数はここ数年の間に急激に拡大している。研究開発を含むプロジェクト／プログラム／政策や、研究開発を実施する機関／組織／人材などを対象とする評価においても、それぞれその事情は同様である。

しかし、我が国においては、研究開発評価の専門家は極めて限られているのが実情である。研究評価の研究者数は極めて限定的であり、また、実務的専門家の数も極めて限られている。敢えて直截に表現すると、研究開発評価のいわば“素人”が評価の大部分を担っていると言うべき状況にあるとあってよい。このような状況は一刻も放置するべきではなく、緊急の措置が講ぜられるべきである。言うまでもなく、評価の専門性の養成と普及を放置するならば、評価の質が高まることはほとんど期待できず、結果として評価に対する信頼性が醸成されないまま形式的な評価システムが運用され続けるといふ、極めて憂うべき状況に早晩至ることが危惧される。後述するが、評価の先進国といわれる国々もかつては同様の課題に直面していたのであり、現在相応の厚みをもつ評価人材を抱えるに至るまで、強化の取り組みが持続的になされてきたのである。

人材の育成にはその質と量に応じた時間を要することが特徴である。高等教育機関のカリキュラムやコースを整備して、そこで養成された卒業生を待ち受けるという対応だけでは必要な人材を供給する時間的なゆとりがあまりにもない。もちろん、そのようなコースの整備も急いで行うべきであるが、そのアウトプットを待ち受けるだけでは当面の評価人材のニーズに対応できない。緊急措置としては、第一に、評価に携わる人材に対して、短期集中的な研修コースを用意し、その履修経験を通じてスキルアップを図ることが考えられる。第二には、研究開発の専門家（研究開発者）に対し評価スキルの修得機会を提供して、彼らを評価人材に転換し、その研究開発のバックグラウンドを活かした評価人材として評価現場で活用する。このような転換コースを整備するということが考えられる。

表 1-1 我が国に必要な研究開発評価人材の養成コースの枠組みと特徴

区 分	目 的	特 徴
高等教育コース	・本格的な評価専門家の養成	・高等教育機関の教育課程として実施 ・普及に時間を要する
短期研修コース	・行政内部等の評価人材の能力向上	・OJTの延長線上にある研修コース ・ポストについて、そのポストで必要とされる能力やスキルの向上を目指す ・短期間に実務者の実践的能力の育成を図る ・高等教育機関が付設することもある
短期転換コース	・本格的な研究開発の実務経験を活かした研究開発評価者の養成	・キャリアパスの転換を目的とした研修コース ・主として研究開発者の評価専門家への転換を目指す ・行政内部等で不足している研究開発経験者の補填

本調査報告書は、このような問題意識のもとで、緊急に整備すべき短期研修コース、及び転換コース、そしてまた、高等教育機関において整備すべき専門家養成コースについて、それぞれの具体的なあり方に関して調査検討し、その成果をまとめたものである。なお、評価に関連する人材をどのように育成供給して「受け入れて用いる」のか、という問題の性格を踏まえて、ここでは人材の「受用」という表現を用いる。

1. 1 人材受用側の全体像

研究開発評価に関わる人材に期待される専門的機能は、主に次の三つの局面にわたっている。

第一は、評価システムを設計し運用するマネジメントの機能であり、第二は、評価過程において評価それ自体を担う機能であり、第三は、評価の前段として、関連する評価対象や状況等についてのデータや資料を集め、それらを分析し、評価のための全体像を明らかにする評価支援機能である。

表 1-2 評価関連機能と担当者(機関)

評価関連機能	機能の担い手(典型)	担い手の所属機関(事例)
評価マネジメント	行政官	行政、ファンディング機関
評価自体	パネルメンバー	研究機関
評価支援	調査分析エキスパート	シンクタンク

これらの専門的機能を担う評価人材は、次のような機関に分散配置されている。

省庁の評価を担う内部部局、ファンディング等の資源配分に携わる中間機関、研究開発自体を実施する研究開発機関、そして、独立に評価を担う独立評価機関、民間にあって評価を支援する調査分析等に携わるシンクタンク、さらには、大学等における評価研究組織や評価研究関連講座等がある。

これらの機関は多くの場合、三つの機能的側面のいずれかひとつを担うというよりも、複数の機能を担う部署を抱えていることが多い。省庁の内部を見ると、評価の統括支援担当課の他に、政策形成とその政策の実施を担当する多くの評価実施担当課、及び評価結果を踏まえて意思決定を行う査定担当部署がある。資金配分機関の場合も、意思決定部署の他に、評価統括支援担当部署と評価実施担当部署から成っているのが普通である。

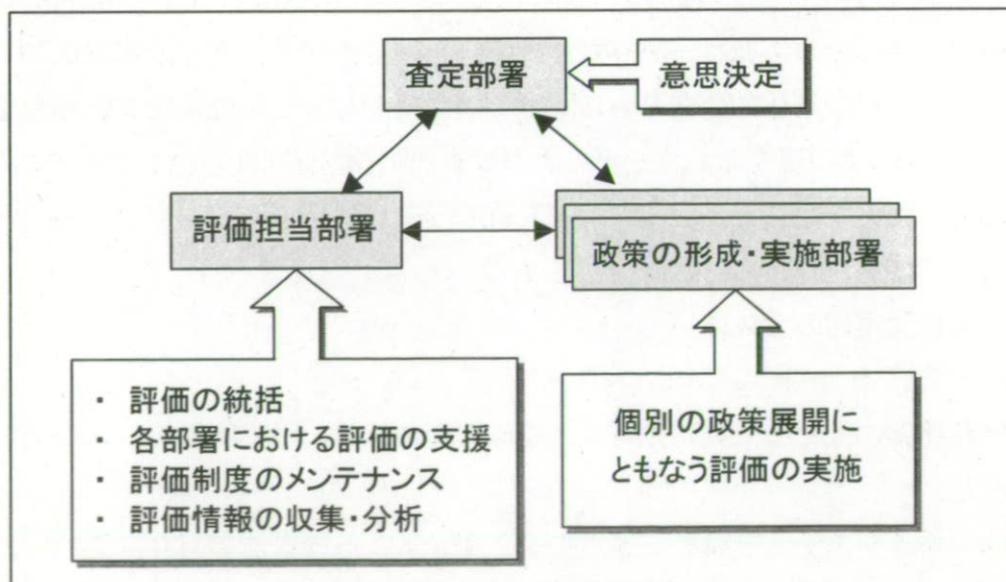


図 1-1 評価関連部署と役割分担

例えば、NEDOにおいては、プロジェクト全体の中間・事後に関わる外部評価の運営をするのは技術評価課（評価統括担当部署）であり、最近の改革の中で実際に新規プロジェクトの事前評価の運営は企画調整部から実施部門（評価実施担当部署）に移管されつつある。しかし現在多くの新規プロジェクトに関しては、評価結果を新規のプロジェクト運営に反映すべく企画立案をする企画調整部と最終意思決定を行う理事会（意思決定部署）との分掌が進んでいない状態であるということになる。

省庁及び資金配分機関の評価統括支援担当課は、評価システム全体を維持管理し、また、評価結果を集約、分析する役割を担っている。またある場合には、評価担当部署それぞれの上位機関が評価を実施することもある。研究開発機関の場合にも、評価の統括と支援を担当する部署の他に、個別の評価を実質的に実施している部署が、機関の研究部門全体にわたって個々に存在していると考えられるべきであろう。大学評価機構のような独立評価機関においても、その評価システムを整備し発展させる役割を担う評価統括支援担当部署と、実際の大学評価の実務に携わる評価実施部署から成っている。これに対し、評価の調査分析支援機能は、シンクタンクや大学などに集積されているべきものである。

1. 2 人材養成側の全体像

本来、評価人材は、その保有すべき専門性と実務能力に鑑みれば、大学等の研究教育機関から育成・供給され、また、シンクタンク等の評価支援機関や行政機関等の評価実施機関の内部に蓄積され、そしてまた、民間企業や国立研究機関等の研究機関の内部にも集積されているべきであろう。実際、評価に取り組み始めた歴史がある程度古い英国や米国をはじめ、評価先進国においては、このような機関等に評価研究者及び実務的専門家が集積されている。

そして更に重要なことは、行政等の評価実施機関の内部で OJT のメカニズムを通じて評価の実務的専門家の拡大再生産が行われてきていることである。また、研究開発の本格的な経験者が不足している行政機関等に対し研究開発経験者を導入するために、高等教育機関以外に専用の研究者転換コースを用意している国もある。

しかしながら、我が国には、明示的な評価システムの歴史が浅く評価専門人材のニーズが小さかったこともあり、高等教育機関に評価人材を養成するコースが設置されてこなかったために、評価人材のストックはほとんど無く、他の専門分野から転向した少数の研究者や実務的専門家が散在しているのみである。また、大学の専門的コー

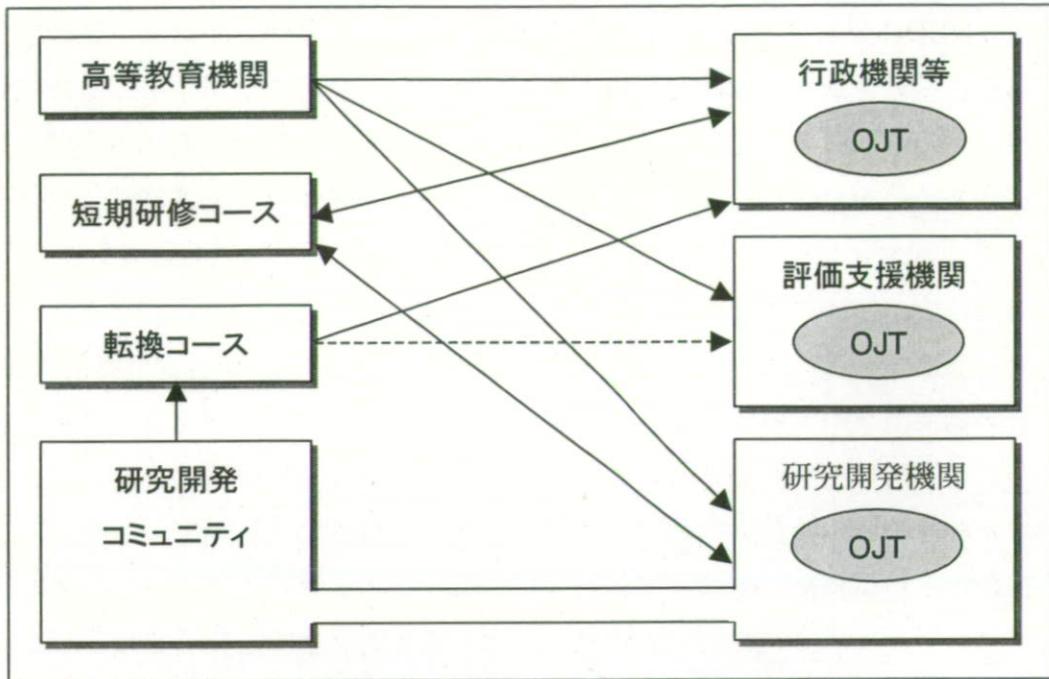


図 1-2 評価先進国における評価人材の養成メカニズム

我が国ではいずれのメカニズムも極めて弱い

スの他に、研修コースや転換コースも少数の民間企業向けのコースを除けば全く整備されてこなかった。この背景には、本格的な評価を必要とせず先行モデルに追随すれば良かったキャッチアップ体制からの脱却が行政機関において遅れているという構造的な要因があると考えられる。民間企業においても、80年代を通じてヒューマンサイド・マネジメントの側にマネジメントシステムの重点が置かれ、評価を通じたサイエンティフィック・マネジメントへの取り組みが重視されてこなかった。

それ以前には我が国では、60年代末から本格的な自由化時代を迎えて研究開発の自主的な推進という新たな課題認識があり、研究開発マネジメント先行国のサイエンティフィック・マネジメントのツールを盛んに導入しようとした時期があった。しかし、オイルショックでこのような試行や研究が頓挫して以降、本格的な評価に取り組む体制が我が国においては構築されることなく、長期間経過した。研究開発評価の日本語の文献を辿ると、70年代初頭までたちまち遡ることからも、この辺の事情を理解できるであろう。この間に浸透した我が国のヒューマンサイド・マネジメントは、いわば「直感的で定性的な評価」を内包したシステムであり、評価が全く行われなかったということではない。むしろ、我が国の評価の特徴として、評価者と被評価者が一体となってその組織が目指す共通の目標に至るための検討を両方の立場から一体として行

うという、「内在的」な評価システムがこの間展開され発展してきたと言える。

一方、英文の文献を辿るならば、例えば、この分野の専門雑誌である“Research・Technology Management”（Industrial Research Institute 発行）所載の文献を辿ると、evaluation をキーワードとする論文は 70 年代の中頃までではほぼ出尽くし、80 年代になると殆どそのようなキーワードを含む論文は出現していない。その代わりに、メトリックスとかメジャメント等の文献が現れ始め、評価を具体的に進める際に必要となる、より困難な問題に研究や考察の焦点が移ってきたことを示している。あるいは、evaluation という用語を用いるにしても、戦略や経済的なインパクト等の評価というように、さらに困難な領域へ評価をめぐる関心が移動していることが伺える。言い換えれば、この間評価実践の体系が深められてきたことを意味する。その一方で、70 年代の前半までもてはやされていた（我が国もかつて盛んに導入しようとしていた）数量的な客観化を目指した評価体系は欧米において質的に深められ、表面上あるいは形式的にはむしろ我が国で展開してきた定性的な方式に置き換えられてきたかに見える。すなわち、ヒューマン・リレーションズ・マネジメントを重視するマネジメント体系が欧米で確立されていく中で、質的に深められた数量的な把握を参考資料として用いると共に、ヒューマンサイドの質を深く捉えるために定性的な判断を重視していく評価体制が欧米で確立されてきたことになる。

このように我が国と先行国との差異は大きい。すなわち、先行国で実施されているメジャメントやメトリックスの有無は、評価の明晰性や信頼性の獲得において大きな差異をもたらすことになった。さらに、メジャメントやメトリックスにおいて、広いパースペクティブのもとで対象全体を捉えようとする努力や、単なる過去のトレンドではなく、将来を見据えた分析やそのためのメジャメントやメトリックスという、より高度な方向にそれらを発展させてもきた。そのため、形式的には定性的な評価を中心とするようになったとはいえ、その質において格段の差をもたらすことになったと言わざるを得ない。

このような専門性に基づく評価を行うためには、専門的な人材が不可欠であり、その養成システムを整備するための施策が急いで展開されなくてはならない。

1. 3 必要とされる評価の教育・研修システムの全体像

上記のような我が国の現状を踏まえ、その状況に適合した教育研修システムを以下に構想する。その内容は三つの側面から成り、まず第一に、いかなる能力を獲得すべ

きか、何を修得すべきかであり、第二には、どのような評価機能を担った評価者を養成すべきかであり、また第三番目には、教育研修システムの構成それ自体に関してである。つまり、評価に関わるコンテンツとアクターとシステムのすべての側面から構想されなくてはならない。

表 1-3 評価人材養成システムを構想するための3つの視点

1. いかなる能力を獲得すべきか (コンテンツ)
 - ・ 評価法だけではなくその基盤となる知識やスキルも
2. どのような評価機能を担う評価者を養成すべきか (アクター)
 - ・ 評価の為のマネジメント人材
 - ・ 評価の為の調査分析エキスパート、
特に社会経済性分析エキスパート
 - ・ 評価そのものを行うレビューア、特に社会経済性評価レビューア
3. 教育研修システムの構成はどうあるべきか (システム)
 - ・ 高等教育機関で行う、体系的な専門家養成コース
 - ・ OJTを補完して、評価マネジメント人材や調査分析エキスパートを
養成するための短期研修コース
 - ・ 研究開発者を評価人材に転換するための短期転換コース

なお、我が国では、高等教育のカリキュラムの中で、評価の背景となるような種類の講義科目がほとんど設定されてこなかった。そのため、高等教育の教育内容を活かして評価の特殊性のみを修得すれば、それで専門的な実務者として通用するという背景的な知識をもつ人材は、それほど多くない。したがって、教育研修の内容としては、評価のみに局在化させるべきではなく、その背景までを含めておく必要がある。

また、評価関係者としては、まず第一に必要な人材は、評価のマネジメントを担う人材であり、省庁の内部と資金配分機関等の中間組織の内部に必要となる人材である。評価のマネジメント人材は、被評価者の持つ評価に関わる様々な疑問に答え、信頼関係を構築しつつ、評価システムの設計や評価委員（評価パネルメンバー）の人選等を行い、公正で中立的な評価活動の運営を担っていかなければならない。第二には、評価のより困難な領域、たとえば社会経済性やインパクト等を分析する専門家である。このような専門家は、本来ならばシンクタンクに集積されているべきであるが、我が

国にはそのような人材が極めて不足しているのが現状である。受け皿であるシンクタンク機関の育成策も併せて検討すべき課題である。また、第三の категорияとしては、評価そのものを担当する評価者である。

とりわけ、社会経済性評価のエキスパートは、我が国では殆ど存在しない。ベンチャー・キャピタリストが極めて手薄であることが、何よりこの間の事情を物語っている。一方、科学技術の質的側面を評価する、科学技術ピアレビューについては、我が国においても長い歴史があり、例えば旧文部省の科学研究費の配分に携わるピアレビューを経験した層は、ある程度存在している。むしろ過去の研究開発評価の運用においては、科学技術ピアレビューが科学技術の質のみではなく、その社会経済性の評価であるとか、さらにはまた、評価システム全体についてのアドバイス等を担ってきた側面もある。問題なのは、その質であり、そもそも科学技術の研究者に社会経済性評価やその分析や、あるいは、それらを含む妥当な評価システムを設計することを期待することの方が無理であったと言わざるを得ない。また、科学技術の社会経済性評価というのは、旧い社会科学の分野のフレームワークでは捉えきれない、極めて新しい発想が必要とされる学際的新領域であることは意識されるべきである。

ともあれ、評価マネジメント人材、社会経済性分析エキスパート、社会経済性評価レビューの3種類の人材を改めて養成しなくてはならない。また、科学技術ピアレビューについては、その本来あるべき姿にそれを戻し、そしてまた、そのピアレビューとしての質を高めその能力を発揮する取り組みが行われなくてはならない。この導入教育は評価マネジメント人材の役割でもある。

このような内容と人材を養成する研修システムとしては、1) 評価マネジメント人材や、社会経済性エキスパートを擁しているべき機関の人材に対する研修コースと、2) 大学や研究機関等に在籍する研究者等をこの分野に新たに導き入れるための転換コースと、3) 高等教育機関で養成する、より体系的な教育を受けた専門家養成コース、の3種類が考えられる。このような内容と人材の種類とコースの具体的な姿に関して以下に詳しく述べよう。

第2章 何を修得すべきか

研究開発評価の実務には、一定の大きな枠組みがあり、これを深く理解しその枠組みを踏まえる必要があるものの、具体的な内容までをマニュアル化することは殆ど不可能であり、また意味がない。むしろ、評価対象と状況に適合した評価システムを独自に設計し、また、その運用に当たっても極めて柔軟に運用することが望まれる。実務的にいえば、研究評価は全て応用問題であり、それを目的とその場の環境と文脈に合わせて解決する必要がある。このような応用的な問題に対処するための能力としては、ワンセットから成る多様な実務的スキルの他に、それらを使いこなすための基盤的能力が必要である。また、研究評価に限定した領域の中のみの実務的能力では不十分であり、その背景にある評価論一般、技術経営論一般、そして、研究開発組織、あるいは評価組織を扱うための組織現象に対する一般的なスキルや能力がさらに必要となる。

表 2-1 どのような能力を修得すべきか

<p>研究評価は全て「応用問題」である。目的、文脈、環境に合わせて評価を適切に行うためには…</p> <ol style="list-style-type: none">1. 評価のための実務的スキル<ul style="list-style-type: none">・ ワンセットの実務的方法論2. 手法を使いこなすための基盤的能力<ul style="list-style-type: none">・ 実務的方法論を使いこなすための基盤的能力3. マネジメント全体を見通す広い視野<ul style="list-style-type: none">・ 評価をマネジメント全体の中に位置づけるための能力
--

さらに評価は、研究開発パイプラインのステージを区切る局面でその都度必要とされる能力であることを考えれば、単に課題の選択のフェーズだけではなく、その課題の妥当性が根拠づけられる計画や、計画のバックボーンとなる戦略等を形成することと一体となっている。そのため、課題の選択に関する評価のみを切り離して評価を適切に行うことは困難である。また、計画や戦略の形成に際しても、それらを選択する

ための評価の局面も存在している。

このように、評価は意思決定問題の前段を構成し、思考世界の判断を司ることになる。したがって、実務的な評価を行うに際しても、このような広がりの中で、思考世界を操作する基盤的能力が必要となる。

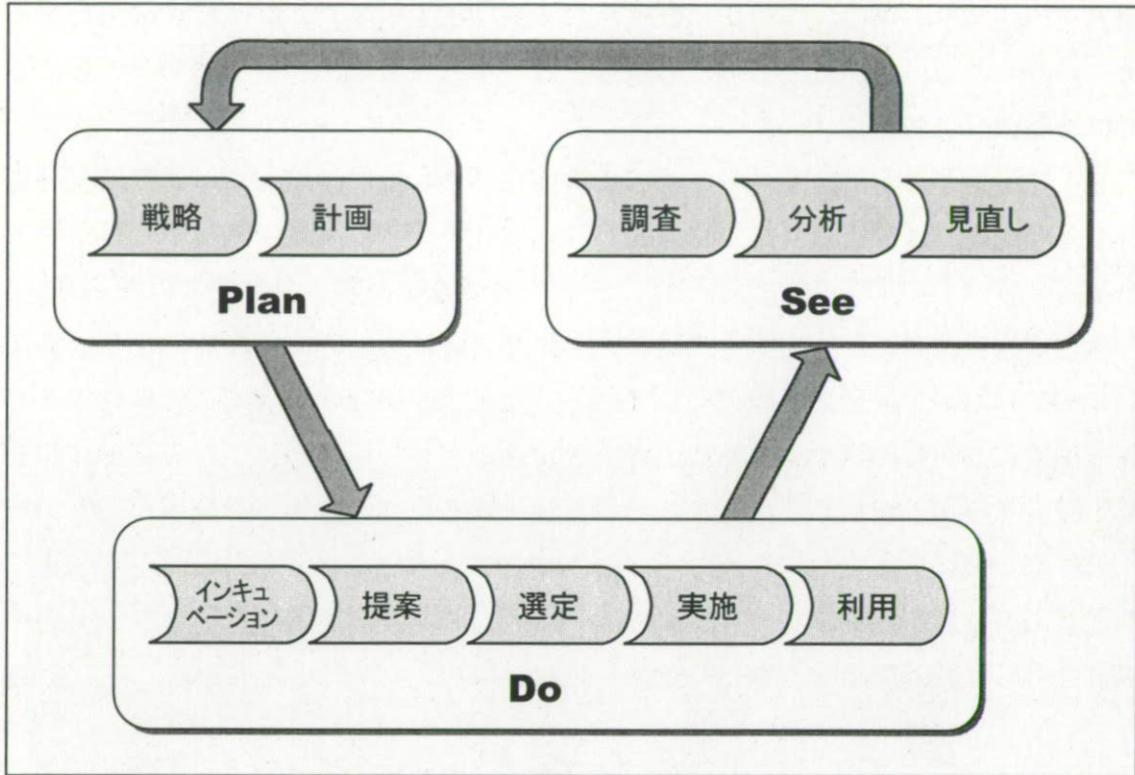


図 2-1 研究開発パイプラインと評価の局面

Plan-Do-See のフェーズ境界だけではなく、各フェーズ内部のサブフェーズの境界においても何らかの「評価」が必要となる。

第3章 カリキュラム内容の概要

前章で述べた能力を獲得するためのカリキュラム内容については、別の報告書『研究開発プロジェクト等の評価手法に関する調査』（2001年3月、財団法人政策科学研究所）に詳しく述べた。ここではその概要に触れるにとどめ、より具体的な内容については同報告書に譲ることにする。

表3-1に示したように、カリキュラムの修得目標としては、第一に、研究開発評価の推進体制の全体像が把握され、そして第二に、評価の諸局面についての理解が深められることが必要である。また第三に、評価の方法論が把握されなくてはならない。その際、研究開発評価に特化することは、しばしば評価の捉え方を硬直的にし、あまり生産的ではない。評価論全般について踏まえた上で、研究開発評価の特殊性を考慮した方法論について理解を深めることが必要である。さらに第四に、評価推進体制を策定し設計できることが必要で、またそれを運用するための枠組みの理解が必要である。その際、評価局面として重要なポイントが押さえられる必要がある。

例えば図3-1に示した、プログラムの下で展開される「従属的プロジェクト」の事前評価という比較的簡単な対象から始め、次に、「プログラム」や「制度」や「施策」

表3-1 カリキュラム内容の例

1. 評価の推進体制	4. 評価推進体制の策定
2. 評価の諸局面	(1) 従属プロジェクトの事前評価
(1) 評価対象の階層性	(2) 機関の途上評価
(2) 評価のフェーズ	(3) プログラムの社会経済性途上 および追跡評価
(3) 評価の内容	(4) 独立プロジェクトの事前評価
(4) 成果と実績の評価	(5) 制度、施策、政策の事前評価
(5) 使命、目的とマネジメントの評価	
3. 評価の方法	5. 事例分析と演習
(1) 評価論	
(2) 研究開発評価の方法論	

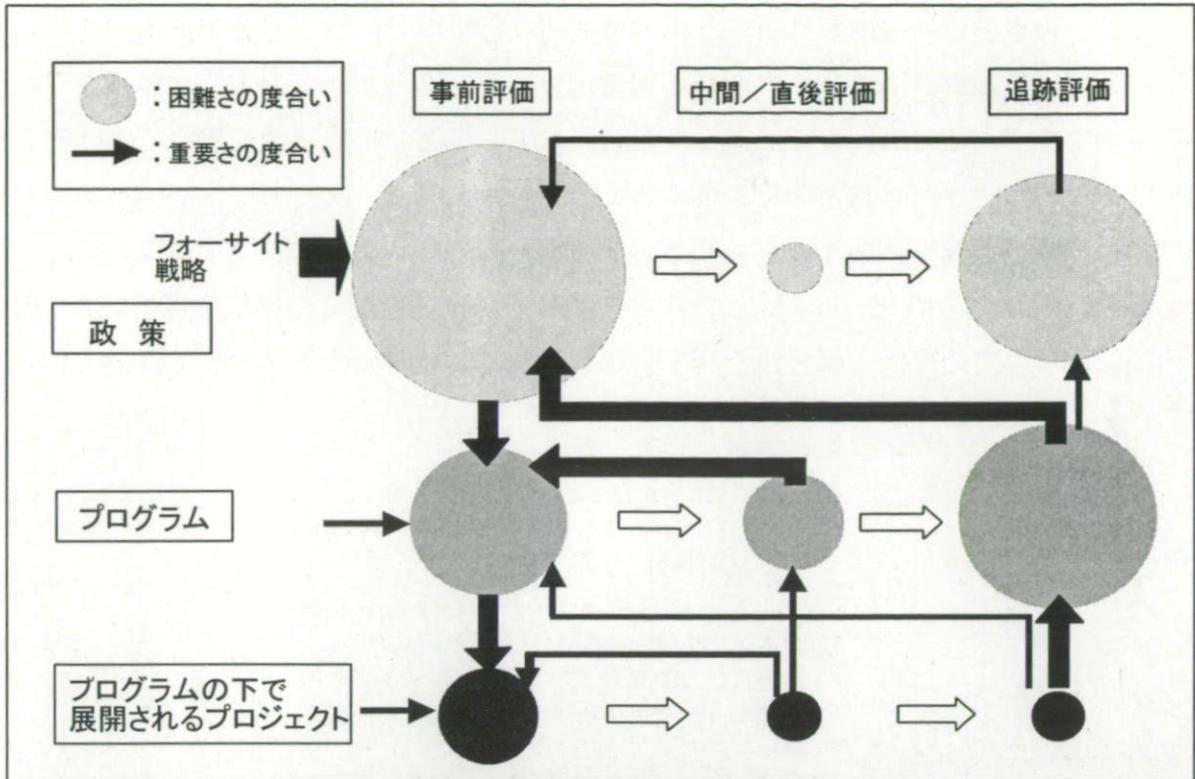


図3-1 評価の階層性と評価活動の困難さ(1)
 円が大きい程評価活動が困難なことを示す

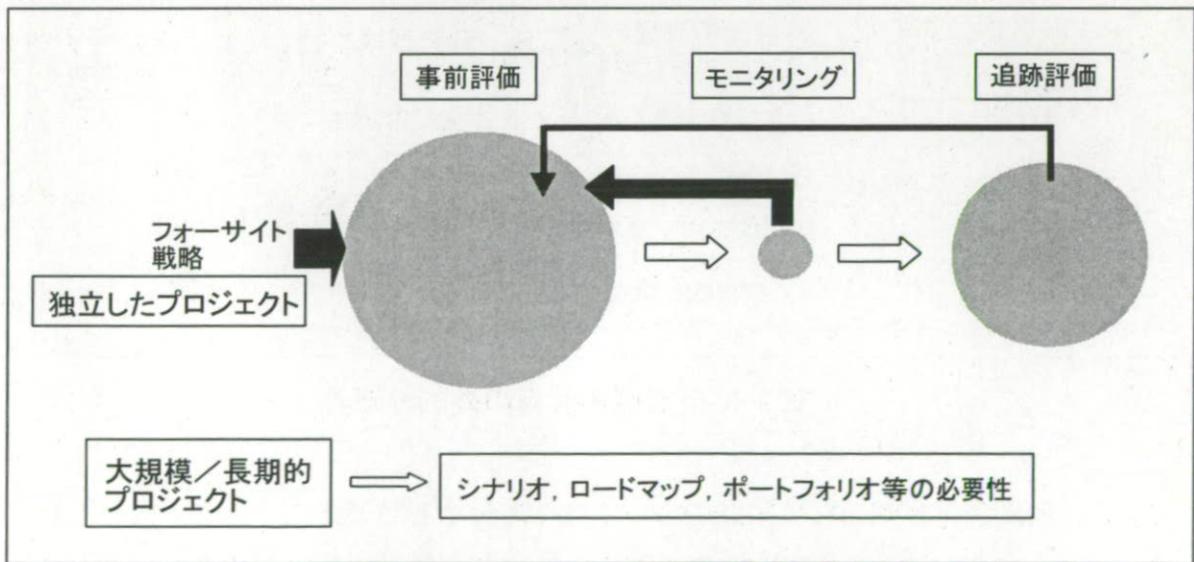


図3-2 評価の階層性と評価活動の困難さ(2)

の社会経済的インパクトを含めた途上評価や追跡評価の方法が理解され、これらを設計、策定、運用できることが必要である。さらには、図3-2に示した、将来に対する先見的な分析を踏まえた評価が必要となる大型の「独立プロジェクト」の事前評価や、

「政策」の事前評価等に取り組むための導入的な枠組みの理解が必要である。これらはいわば研究開発評価の体系に関わる理解についてであるが、そのような体系を理解し使いこなすためには、さらに進んで第五として、事例分析や演習を重ねて実践的な理解に深めることが必須となる。

その際、研究開発評価に限定されたことのみを研修することは賢明ではない。先にも述べたように（図 3-3 を参照）、方法論の基盤や科学技術経営の枠組み等の、より広い領域についての理解を深め、それらの知識が研究開発評価の前提として修得されることが必要である。

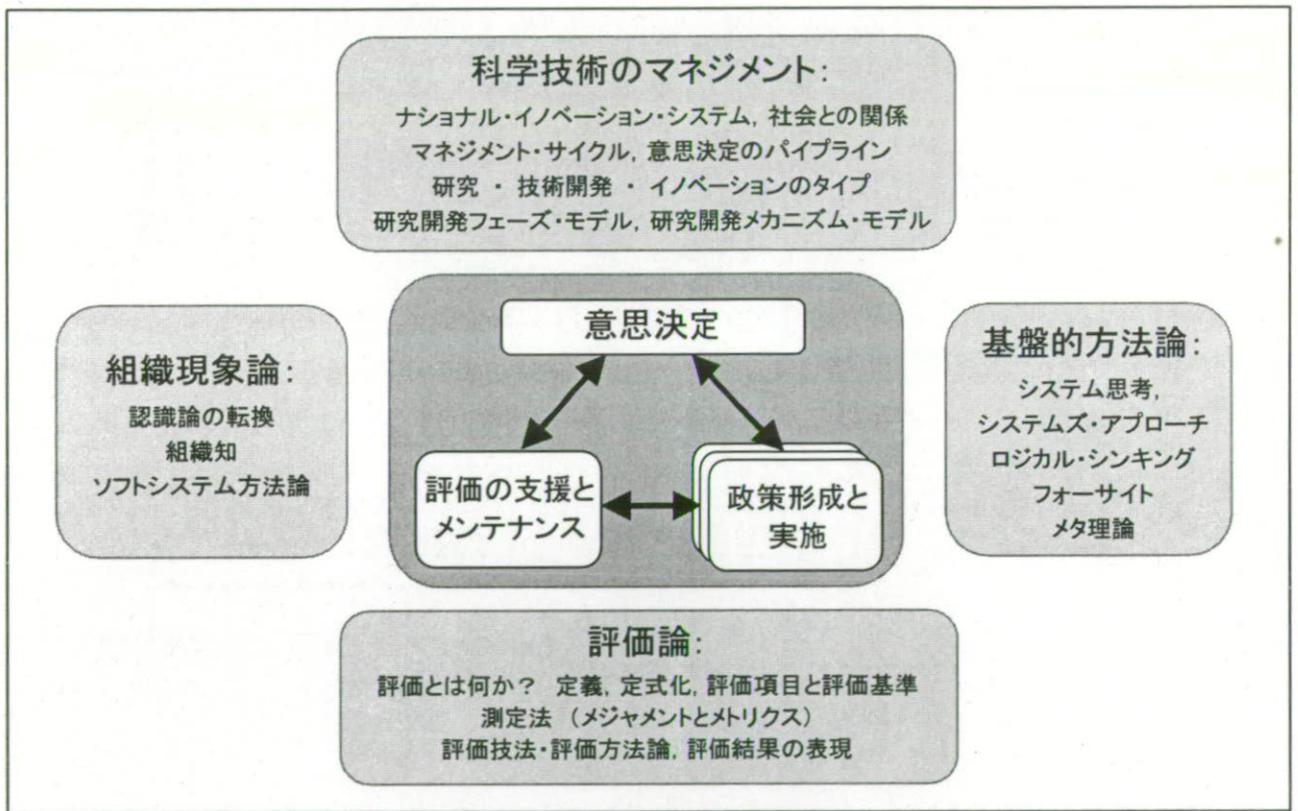


図3-3 研究開発評価の外的枠組み

また、研究開発評価固有の知識について深く理解する必要がある。この内容としては、図 3-4 にまとめたように、評価の理念に始まり、評価対象に関する1セットの知識、更に評価のためのデータの収集を含む評価法の体系、および評価に関わる人材の広がり等に関してである。

なお、このような入門的な研修内容を修得した後、さらに発展的な研修コースに進むことが必要である。世界的に見て、研究開発評価に関わる知識は現在急速に進展し

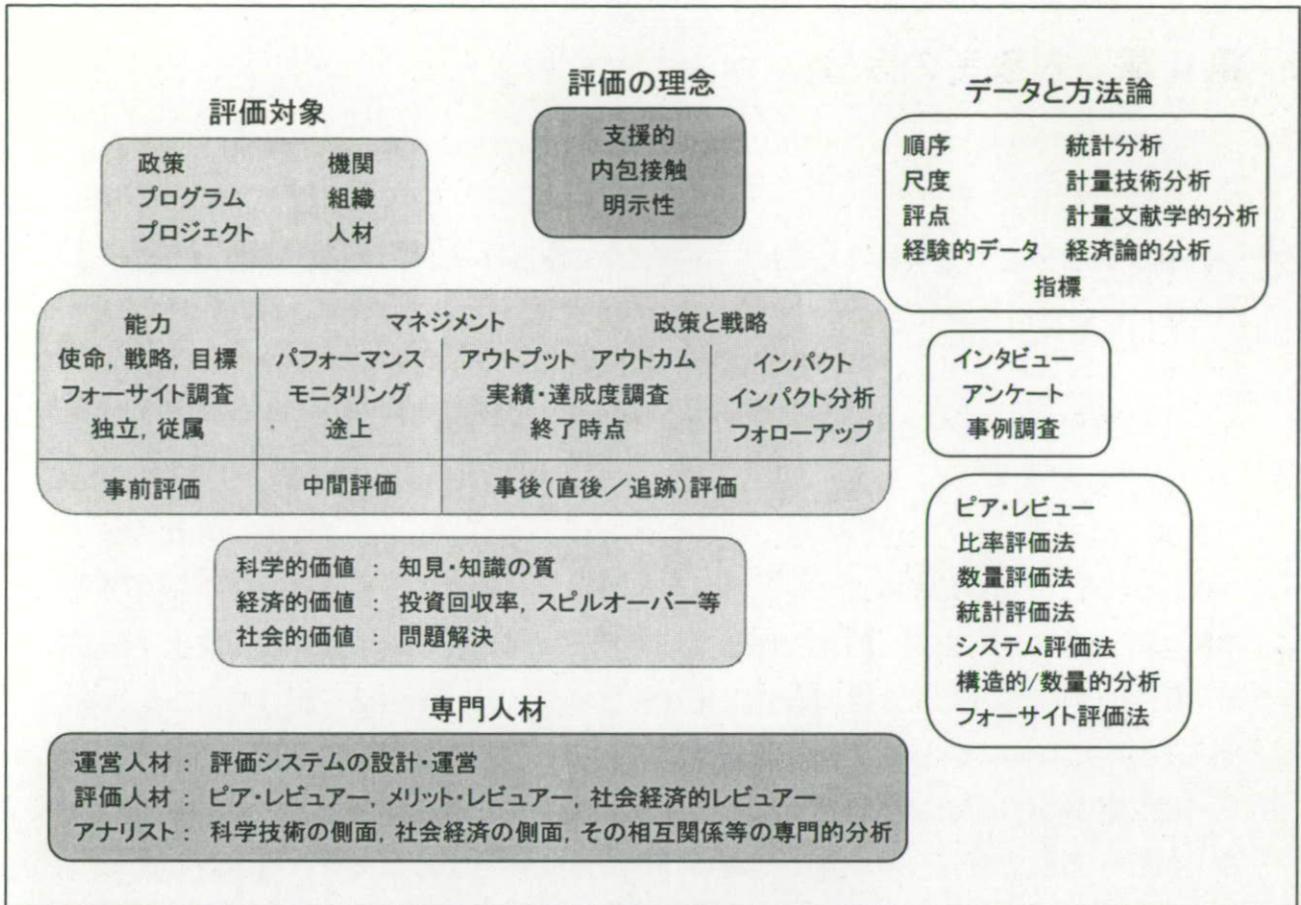


図3-4 研究開発評価の内部構造

ており、それらの知識は欧米の評価先進国間のネットワークで共有されていたり、交流のための国際的な研究集会が持たれたりしている。したがって、海外の先進的な研究評価事例についても理解を深めることが必要であり、この機会を提供する「発展的な研修コース」を整備することが更に必要である。

第4章 対象者の区分

評価システムを担う人材としては、評価マネジメント人材、社会経済性分析エキスパート、社会経済性評価レビューア、科学技術ピアレビューアの4種類が重要である。

4.1 評価マネジメント人材

評価マネジメント人材は、大きく2種類に分類することが出来る。第一のカテゴリーは、いわば研究開発マネジメントの専門家であり、省庁や資源配分機関や研究開発機関の評価システムを統合的に管理運営し、各所で実施される個別評価を支援する部署に配置されており、より広い研究開発評価マネジメントの知識や能力を備えた人材である。ここでは、評価マネジメントのコア人材と呼ぶことにしよう。第二のカテゴリーは、評価実施部署に配置され、その評価を担当する研究開発評価マネジャーである。担当する研究開発課題の評価に関わりその運営に当たる人材である。前者に比べ、評価対象がかなり限定され、担当する対象課題のみに取り組む限定的な立場の評価マネジャーである。人数的にはこの種の評価マネジャーの方がかなり多い。

第1の専門的なコアマネジャーは、入門的な評価マネジメントコースでの研修の他に、さらに進んで発展的な評価マネジメントコースで研修することが望ましい。また第2の、限定された対象の評価マネジメントに携わる評価マネジャーであっても、入門的な評価マネジメントコースを研修することが望まれる。

4.2 社会経済性分析エキスパート

科学技術以外の状況をここでは、「社会経済性」と呼ぶことにする。研究開発課題はその目的や評価の視点から3種類に区分できる。第一は、科学技術の内部に存在する価値を追求する課題であり、第二は、科学技術の外にある課題のうち、経済的な課題を追求するものである。第三は、それ以外に分類されるもので、社会性を追求する課題に代表させることができる。

科学技術の価値に関しては、その科学技術の対象である専門分野の研究者は、その専門性を生かして、かなり深い判断をすることが出来る。研究者の属する専門学会自

体が、当該の専門分野の研究者同士の相互評価（同僚評価：ピア・レビュー）体制により、研究成果を歴史的に形成・蓄積することを基盤としており、研究者には馴染みのある体制である。

これに対して、経済性、あるいは社会性に関しては、そのような専門家は極めて限られている。特に、実務的な領域でそれらの専門性を備えた専門家は、社会の中に広く薄くしか分布していない。また、その社会経済性の具体的状況を分析するためには、関係者に対するインタビューや質問紙による情報収集等を通じたオリジナルなデータの集積や、あるいは、関連している局面の統計資料を手がかりとした分析から得られる状況認識等を深めていくことが必要であり、そのためには時間をかけた分析が必須となる。科学技術についても、もし、その分野を専門としない専門家が取り組むとすれば、同じような困難さがあることになるが、幸いにして、サイエンス&テクノロジー・コミュニティはそれなりに組織化されているために、その対象としている専門性を深く理解できる専門家を比較的容易に見つけ出し、その専門家による判断を仰ぐことが可能となっている。しかしながら、社会経済性のそのような実務的側面に関しての専門家は、一般的には存在しない。つまり、課題と状況に合わせたオリジナルな分析がここでは必要となる。このような内容を担える実務的専門家が改めて養成される必要性がある所以である。

社会経済性分析のエキスパートは、入門的な研修コースの他に発展的な研修コースで研鑽する必要がある。特に、外国の研究機関で深められているそのような分析作業を事例として辿ることは、多様な評価方法をどのように組み合わせて対象と状況に適した評価システムに組み上げているかを具体的に理解することが出来、そのような事例を多く経験する中で、分析のエキスパートとしての能力が開発されることが期待される。さらには、海外の研究者によって行われている先進的な研究成果を学ぶことによって、新しい方法論の観点を付け加えることが出来るので、このような研修もまた必要である。社会経済性分析のエキスパートを養成することは、容易なことではない。

4. 3 社会経済性評価レビューア

社会経済性評価レビューアは、とくにプログラムに従属するプロジェクトの事前評価において、そのプロジェクトが社会経済的効果を目指している場合には必須の人材となる。この場合、いわゆる従属的プロジェクトの評価パネルとして、科学技術ピアレビューアによるパネルの他に、経済社会性評価のレビューパネルを構成す

ることになる。あるいは、この両者をひとつのパネルに統合して評価パネルを構成することもよく行われている。なお、このような従属的プロジェクトの社会経済的な側面の評価を限定的なかたちで進めてよいと考えられる場合がある。第一は、プロジェクトの投資規模がそれ程大きくはない場合である。第二は、大きな目標としてのプログラムが設定された中でのプロジェクトの展開であるために、限定的な社会経済的な側面を評価すれば十分となるようなケースである。

機関評価においても社会経済性を評価する必要がある場合は多い。しかし、機関評価の場合には、内部評価によって事前に社会経済性評価に耐えるデータが研究者によって収集されていることが期待される。あるいは、研究者の側で収集することが十分でない場合には、機関評価を支援する社会経済性分析エキスパートによって、それらのデータが整えられていることが期待される。

独立した大型の長期的な研究開発プロジェクトの場合には、その社会経済的効果を評価パネルのみによって判断することは殆ど不可能である。手続きのにも、その前段として社会経済性分析が深められている必要があり、それらのデータに基づいた判断を評価パネルが行うことになる。

また、プログラムや制度・施策等の途上評価や、あるいは追跡評価はもとより、政策の事前評価等においても、当然ながら、社会経済的な側面に関しては、独自に深い分析が事前になされていることが必要であり、それらのデータを参考資料として整備・提供することを前提に、評価パネルが形成されることとなる。

このように、社会経済性を対象とした評価パネルの場合に、社会経済性分析のデータが存在する場合とそうでない場合とが想定される。社会経済性分析のデータが事前に整備されている場合のパネリストとしては、社会経済性分析を多く経験したこの分野のシニアエキスパートが当たるべきであろう。しかしこのような人材は極めて限られているのが実情で、前述の社会経済分析のエキスパートを急ぎ育成し、この任に当たらせるべきである。

また、社会経済性分析を使命とした評価パネルのパネリストの場合には、いわゆるエキスパートパネルを形成して評価にあたることになるが、そのエキスパートは、シンクタンクや行政内部で蓄積されているこの方面の社会経済性の分析の専門家が当たるべきである。いずれにしても、そのための人材の養成・確保の問題を考慮しておく必要があり、これもまた、前述の社会経済性分析のエキスパートの養成コースに依存することになる。

前述のように、我が国では現在、この分野の専門家は皆無に等しく、代わりに、年

長で経験をもつ科学技術人材を登用したセニョリティーパネル、著名な研究者・技術者によるエminentパネルを構成して評価を委嘱しているのが実態である。しかし、彼らに対してかつて社会経済性評価固有の能力開発機会が与えられていたわけではなく、評価が実質的に深く成されているとは言い難い。むしろ、社会経済性評価の質を著しく損ねてきたと言うべきであろう。また、彼らに研修機会を与えたとしても、自己流の権威の殻を打ち破ることは容易ではなく、ほとんど進歩は期待できない。この意味からも社会経済性分析のエキスパートの育成が強く望まれる。なお、その際評価者への転換の意思をもつ中堅ないしシニアリサーチャーから評価者への転換には大いに期待がもてる。転換研修コースの充実も重要な課題である。

一方、分析データを前提としない比較的単純な枠組みの中で行われる従属的プロジェクト評価の社会経済性パネルのパネリストは、広く社会の中に存在していて、そのような専門家達の直観的な能力に判断を依存することになると思われる。例えば、ベンチャー・キャピタリストをはじめとするアーリーステージの企業を立ち上げるためのサポーター・システムに関与する専門家、企業の事業部の企画部門に携わる専門家、あるいは地域で企業のコーディネーションに携わる専門家等が想定される。このような局面でも、従来、我が国でしばしば経験してきた失敗は、企業が社会経済的な活動をしているからといって、社会経済性評価のレビューアとして、企業の研究所の企画部門の担当者を採用することである。民間企業であっても研究所内の人材による評価は、改めて一定の能力開発と方向づけを行わないかぎり、いわゆるシーズ側から、つまり技術の側から社会経済性を押し量るシーズプッシュ・タイプの評価になりがちであり、信頼できる意見には成りがたい。企業の研究開発に関係した部署であっても、事業部であるとか、あるいは営業部門の経験の豊富な担当者であることが望ましい。とはいえ、彼らに対して、評価パネルを構成するために改めて研修コースでの研修を義務づけることは殆ど不可能であろう。

4. 4 科学技術ピアレビューア

科学技術のピアレビューアが評価すべき局面は、科学技術の質に関わる問題である。科学技術のピアレビューアは、評価対象の科学技術的専門分野に通じた研究者や技術者である。言うまでもなく科学技術の専門性は非常に深いものがあり、その専門的な内容を理解できる人は、ほぼ専門を共有する科学技術の専門的研究者、技術者に限定される。科学技術は累積型の学問体系を成していて、深い専門性を理解するためには、

その前段となる基礎的な事項をよく理解している必要があり、まさにそれはその分野の専門家である条件となっている。研究開発評価はどのような課題であるとしても、このような科学技術的な側面がまず第一に評価される必要があり、科学技術ピアレビューはあらゆる研究開発評価の課題に対して必須の存在である。

科学技術ピアレビューが、研究者技術者としての評価者であるならば、評価者としての役割は、いわば専門に対して付加的なものであると理解すべきであろう。その意味で、ピアレビューに課せられる研究開発評価に係る知識は、他の評価専門家たちとは異なり、評価に対する常識的な知見で十分であると考えられる。仮に評価に関わる専門的知識が必要であった場合であっても、研究開発の第一線で活躍する研究者技術者が評価研修コースに参加することを期待するのは殆ど不可能であろう。したがって、逆の見方をすれば、科学技術ピアレビューに期待する評価局面は、科学技術の専門的な質に限定すべきである、と考えられる。しかしながら、そのような場合であっても、評価に関わる基本的な枠組みについての理解は必要である。それは評価を運営するマネジメント人材によって、科学技術ピアレビューに対するガイダンスの形で、評価に際してインプットされるべき知識と捉えるべきである。

我が国においては、評価者が科学技術ピアレビューのみであると考えられるような原初的なステージの時期においては、科学技術のピアレビューが、社会経済的な側面も含めて全ての評価を担当し、さらには、評価システムの設計にまで携わるということが期待された。もちろん科学技術の専門の研究者技術者がそのような評価者に対応できる知識を修得し、その評価に取り組むようなことがある場合も想定される。しかし、そのようにオールマイティの評価者たれんとする志は歓迎するとしても、全てのピアレビューにそのような機能を発揮することを期待することはとうてい不可能であるし適切ではない。むしろ今後は、社会経済性評価レビューのような新たな専門性を備えた評価者を養成することによって、科学技術ピアレビューの役割を限定的にし、またその質をその限定的な領域で高めることを目指すべきである。

このように想定した場合、科学技術ピアレビューに対するガイダンスとしては、評価者として期待されているものが何であるかについて確認することが基本となる。つまり、科学技術の質を専門的に評価するということであり、またある場合には評価対象である特定の狭い領域を科学技術のより広い領域の観点から評価するというように、あくまでも科学技術の質を測ることに科学技術レビューの能力が限定されていることについての認識を促すべきである。ガイダンスでは、このような評価者としての位置付け、期待されている能力、そして評価システムの基本的枠組みの理解がま

ず第一に図られる必要がある。

その次に、価値的側面に関わることになるが、評価のクライテリアを何に取るかということに対して、ガイダンスが成されるべきである。優秀な研究者技術者は、高い成果レベルを目指し、知的フロンティアへの挑戦を彼自身の研究活動の中で日常的に行っていると理解できるが、評価のクライテリアとしては、まず、そのような知的フロンティアへの挑戦の有無、あるいはその程度といった尺度が取られるべきである。また、その具体的な中身は、研究分野の特性に応じて評価者が構築すべきものと考えられる。また、別のクライテリアとしては、プロジェクトの性格に応じて、研究の基盤の広さであるとか、期待される科学技術上のインパクトの大きさ等の評価基準もあり得る。どのクライテリアが適切であるかは、評価マネジメント側からのガイダンスも必要であるが、また、ピアレビュー自身が深く考えるべき内容でもある。そして定められた尺度のもとで、あるいは方向性のもとで、専門的な質が測られなくてはならない。さらに、ガイダンスにおいて提供する不可欠で具体的な知識として、評価制度の規程や評価者としての倫理的な要請等に関するものがある。

ピアレビューは大学や研究機関に広く存在しているが、専門的な深さと同時に、高い見識とか広い視野を備えた研究者が評価者としてより適していると言える。そのような評価者のリストは、評価の実績を通して、また、適宜推薦や公募を含めた様々な拡充の機会を通じて、徐々に充実されて行くべきである。こうして評価システムの試行錯誤の中で妥当な評価を行う研究者たちを見出していくことも、評価マネジメント人材の役割として重要である。もともと、評価のマネジメントにおいては、こうした科学技術側のピアレビューを選抜したうえで、動機づけ、評価目的に沿って備わっている評価能力を十分に引き出すことが肝要である。また、評価過程の情報の開示や、卓抜な評価者を選定する意味でも「評価者の評価」などを実施して、評価者としての適切な緊張感ある役割を果たすことを促すことも有効である。

以上記述してきたように、科学技術ピアレビューの研修コースは、特に設定する必要は無いと思われる。しかし、科学技術ピアレビューに対するガイダンスの内容については、綿密に練り上げられる必要がある。また、研究者技術者の中から、評価やマネジメント一般に興味を持つ人材を発掘し、そのキャリア転換を図る転換コースに誘導することは重要である。また、転換コースを超えて、さらに広い経験を有する研究者技術者を評価マネジメントの責任者としてリクルートすることも考えられる。このような人材の場合には、研修コースに加入するかどうかは別として、マネジメントの一通りの研修ないし履修を期待すべきであろう。

なお、我が国では明示的な評価制度の歴史が浅いこともあり、経験を持った適切なピアレビューアの層が薄く、少数の評価者への負担が増して、「評価疲れ」現象のあることが、先行国と同様に指摘されている。加えて、知的フロンティアの拡大や学術の細分化現象の中で、実質的にピア・レビュー能力を持つ研究者・技術者の層が薄くなり、また、結果的に利害関係をもたない評価人材が少なくなるという問題も生じているという指摘もある。また、歴史的には学会の成立以来、研究者の責務とされた評価活動に対しても、インセンティブや責任をもったシステムが必ずしもうまく機能しなくなったともいわれる。これらは、単に科学技術のピアレビューアの確保という問題にとどまらず、研究開発コミュニティの基盤の健全性に関わる問題として、広く専門の学協会や高等教育機関と連携して実態の把握や対応策の検討を始めるべきである。

第5章 研修コース案

この章で検討する研修コース案は、評価関連機関の関連部署に在籍する職員を対象としている。コースの内容としては、本格的入門コースである「マネジメント・コア・コース」と、発展的なコースに位置付けられる「エキスパート・コース」から成っている。

入門的なコースは、我が国の現状を深く理解した上で設定されるべきであり、このようなコースを海外の実務者や研究者に委ねることは危険であろう。たとえば、現在我が国で実施されている多くの既存研究開発プロジェクトは、目標の設定が明確でなかったり事前評価が十分になされておらず、そのようなプロジェクトを対象にした中間評価や直後評価を現在やらざるを得ない変則的な状況にある。この様な日本の特殊事情をよく理解しその事情を踏まえて臨機応変に評価に対応できる柔軟な発想の養成が研修内容として必要である。

これに対して、発展的なコースはむしろ積極的に海外の事例に学ぶべきであり、海外の実務者や研究者を講師としたコース設定であるべきであろう。特に、大型プロジェクトの事前評価やプロジェクト終了後5年から10年経過した後の追跡評価等については日本に適切な評価事例がないばかりでなく、実施経験者もほとんど見当たらない。必然的に海外での評価事例に学ぶことになり、したがって講師もそのような経験をつんだ海外のエキスパートに依存するのが適切であろう。

表 5-1 研修コース案の概要

-
1. マネジメント・コア・コース(本格的入門コース)
 2. エキスパート・コース(発展コース)
 - (1) 海外評価実践事例に学ぶシリーズ
 - A. 社会経済性分析コース
 - B. 研究開発評価システム・コース
 - C. 研究開発プログラム・コース
 - (2) 海外研究事例に学ぶシリーズ
-

また、発展的コースとしては、実務的な事例に学ぶコースと、研究的な成果に学ぶコースの2種類を区別して用意すべきであろう。その理由としては、研究開発評価に関する研究の歴史をひも解けば明らかなように、圧倒的に多くの研究が非実用的な形態のまま集積していることである。つまり、殆どの研究論文は実務的には採用されがたい適用限定性や展開困難性をもった、研究者からの提案にとどまっている。したがって、実務的に使われている事例を学ぶことを第一に行った上で、実務的な観点から、さらに深めた知識を必要とする局面等に焦点を絞って、その限定された課題に関しての研究者からのコメントや研究開発事例の紹介を得るというステップが有効であろう。海外の研究者の場合であっても、研究的側面からの興味を彼らに任せて披露・紹介してもらうことは、研究開発評価の研究者にとっては有効であっても、我が国の制度・文化・経済社会条件のもとで業務にあたる実務者にとってはそれ程有効であるとは言えない。

しかし一方で、海外の研究者であっても実務的評価に多数関わった経験豊富な研究者も一定程度大学に在籍していることを付け加えておきたい。彼らこそベストの講師たりうる。

5. 1 マネジメント・コア・コース（本格的入門コース）

本格的な入門コースとしてのマネジメント・コア・コースでは、評価対象に相応しい評価システムを設計し運用できる人材の養成をめざす。その内容は、表 3-1 に示したカリキュラム内容のほぼ全域に亘って行われるべきであろう。その主眼とする目的は、日本的な研究開発システムの理解、日本的な研究開発評価システムの実態及びその欠陥や修正の方向等に関する知識、そして実務的な運用スキルの修得である。研究開発評価のマネジメントに携わる人材（コア人材であるかどうかにはよらず）、及び研究開発評価の支援業務に携わる専門家（エキスパート）は、まずこのコースを修得すべきである。

このコースの到達目標は、

- a) 研究開発評価システム全般に関する知識の修得
- b) 研究開発評価システムの外側にある技術経営マネジメントに関わる知識の修得
- c) 研究開発評価のメジャメントとメトリックス及びそれらを統合する方法論についての全体的なスキルの修得

である。このような研修経験を通じて、評価方法のワンセットを理解し、結果とし

て対象に合わせた評価システムの設計と運営が可能になることが期待されている。

カリキュラムの内容から判断し、このシリーズとしては、演習や事例分析を含めて研修期間は5日間程度が妥当と思われる。

表 5-2 マネジメント・コア・コース

- ・ 「本格的」入門コース
- ・ 国内講師による
- ・ 研究開発評価システムの設計や運用を担う「評価マネジャー」の養成をめざす
- ・ コアマネジャーやエキスパートのための入門コースでもある
- ・ 研修期間:5日間

[カリキュラム内容]

- | | |
|------------|-----------------|
| 1. 評価の推進体制 | 4. 評価推進体制の策定と運用 |
| 2. 評価の諸局面 | 5. ケース・スタディと演習 |
| 3. 評価の方法 | |

5. 2 エキスパート・コース（発展コース）

本格的入門コースの修了者を含め、一定の評価実務の経験を積んだ人材に対して、さらに発展コースとしてのエキスパート・コースを用意する必要がある。このコースでは、主として評価のための調査分析を担えるエキスパートの養成をめざす。我が国には残念ながら発展コースの教材として適切な評価事例に該当する先進的な事例はまだ集積されていない。したがって、この領域に関しては、海外の高度な実務者、研究者から指導を仰ぐ必要がある。エキスパート・コースは、「海外の実践事例に学ぶシリーズ」と、「海外の先進的研究に学ぶシリーズ」の2つに分けて設定すべきであろう。

A. 海外実践事例に学ぶシリーズ

・社会経済性分析

我が国の評価人材の育成上、今日、急いで修得すべき発展的な課題としては、高度な知識や複雑な手続きを要するという点で、取り組みにおいてより困難な局面ではあるが、その必要性が高い分野に焦点を絞るべきであろう。まず第一には、社会経済性分析に関わる課題である。社会性及び経済性に関しては、その把握が困難であるばかりでなく、また、これを扱う実務者が我が国において欠乏してもいる。すなわち、その把握を専門とする人材が、専門性の広さ故に通常の養成メカニズムの中では十分には養成することができなかつた分野である。

表 5-3 エキスパート・コース

-
- ・ 発展コース
 - ・ 海外講師による
 - ・ 研究開発評価のための調査分析を担える「評価エキスパート」の養成をめざす
 - ・ 研修期間:3~5日

- (1) 海外実践事例に学ぶシリーズ
 - (2) 海外研究事例に学ぶシリーズ
-

りでなく、また、これを扱う実務者が我が国において欠乏してもいる。すなわち、その把握を専門とする人材が、専門性の広さ故に通常の養成メカニズムの中では十分には養成することができなかつた分野である。

社会経済性に関わる課題をさらに絞るとすれば、2つの課題に収斂するであろう。第一は、大型の研究費を使う、独立型の大型プロジェクトの社会経済性事前評価である。通常大きな予算を長期間に亘って投下するが、そのために、その可能性に関して適切な評価が必要となる。また、もう一つの課題は、プログラムレベルの社会経済性インパクトの途上及び追跡評価である。政策への反映あるいは政策の見直しのために評価が非常に重要なステップとなっている。

このような課題に取り組む専門家は、欧米においても行政内部には少なく、専ら行政の外部のシンクタンク等に分布している。欧米の場合、そのようなシンクタンクは、自らが保有する分析ツールの特性に応じて、それぞれの分析のための得意分野を有し

て実績を重ねている。

研修コースの内容としては、彼等専門家が取り組んだ事例の紹介を中心にするのが望ましい。社会経済性評価に関してはワンセットのツールが存在しているとはいえ、課題に応じてそのいずれをどのような組み合わせで使用するかが重要であるからである。通常、社会経済性評価に関わる分析は、いわばケーススタディとして、事例的に分析が深められるということが一般的であり、その際多面的なツールを動員し、対象に応じた分析の方法を設計して取り組むこととなる。したがって方法論を研修の主題にするのではなく、事例を主題にし、そこでどのようなアプローチが展開されたかを学ぶことが、より重要な研修課題となる。望ましくは、それぞれ特色ある分析ツールを保有するシンクタンクに1シリーズずつの研修コースを担当させ、その得意分野の紹介を受けることであろう。

例えば、独立した大型プロジェクトの社会経済性事前評価に関しては、米国のエネルギー省 DOE の大型プロジェクトを評価している Battelle 研究所がある。また、プログラムレベルでの社会経済性インパクトの途上及び追跡評価に関しては、大型シンクタンクは総じてどこも取り組んでいるが、その中でも SRI International や ADL、あるいは教育分野を中心にした場合では Abt 等がその例として挙げられる。

対象とする事例は、非常に大きな課題であるために、それらの全容を紹介するためには、1シリーズとして5日程度の時間を見込む必要がある。例えば SRI International が担当して現在進めつつある NSF のエンジニアリング・プログラム全体の途上評価においては97年以来、3次に亘る報告書が出されており、現在もその分析は継続中である。その視点は単に NSF プログラムの内部のみを精査するだけでは

表 5-4 社会経済性分析コース

-
- ・ 2大課題: 「独立大型プロジェクトの事前評価」と
「プログラムの途上、追跡評価」

[評価事例]

- ・ Battelle 研究所: DOEの大型プロジェクト評価
 - ・ SRI International: NSFのエンジニアリング関連プログラムの評価
-

なく、他のエンジニアリング・プログラムと比較した場合の NSF によるファンディングのパフォーマンスを論ずる等、広い範囲の分析が行われている。方法面でも、システム的な分析ツールを中心とするが、ビブリオ・メトリックスによる数量的分析や、関連研究者等へのインタビュー調査を中心とした定性的評価など、多様なアプローチが総合されている。このような取り組みの主要部分の紹介を受けるとした場合、事例に携わった複数の実務的専門家を招聘し、それぞれの専門性の深さを紹介してもらうことが有効であろう。

・ R&D 評価システムの設計と運営

R&D 評価システムの設計は、入門コースの知識を修得しこれをベースにすることによってかなりの程度の設計が可能であろうが、さらに状況に合わせて適切な評価システムとして設計・改善していくためには、多くの設計事例や運営の問題点等を学ぶ必要がある。

このような評価システムの高度なレベルでの設計や運営は、その専門家たちを組織して対応することが多い。行政内部に評価マネジメントのための専門性を集積し運用することが困難な場合には、外部に存在する評価マネジメントの専門家たちの支援を受けることになる。

例えば、米国の場合、州政府レベルの科学技術政策の展開に必要な評価業務に関しては、州政府内部にそのような専門家を集積することが困難であるため、NSF からの資金により AAAS の実務担当者が事務局となって専門家を集め、州政府の評価システムに関しての助言を行い、その高度化に取り組むメカニズムができています。このような仕掛けの中で、10 以上の州政府レベルの研究開発評価システムが AAAS によって設計・運営されている。これらの実体は多様であり、そのような多様性を踏まえて事例的に、学ぶことは大いに意味がある。

また、ヨーロッパにおいては民間の専門コンサルティング企業グループであるテクノポリスが支援する評価システムが幾つか存在している。イギリスのテクノポリスの場合には、社会経済性部門を含むプログラムの評価システムの設計やその運営の支援のために、テクノポリスが単なる庶務的な業務以上の寄与を事務部門サイドから行っている。ここで果たしている機能は、運営に対するアドバイスと代行である。テクノポリスは欧州の 6 カ国にいわば独立した拠点支社を展開しているが、そのような拠点が存在しない国に対しても、同じ立場からのサポートを行っている。例えば、イギリスのテクノポリスは、ノルウェーの科学技術政策評価システムに対する設計・運営を

サポートしている。

また、米国の NASA のサイエンス・プログラムの評価活動に対する支援は、中小企業振興政策に基づき大型でないシンクタンクを対象として公募され、そのスキルを備えたシンクタンクが数社選定され、数年間に亘りその運営にあたっている。その一つに InDyne が挙げられるが、その中心的な実務的専門家は、NSF のプログラム・ディレクターを長年務めた高度な実務者であり、NSF からのスピニアウトによっている。このような評価マネジメントに関わる実務的側面の専門家による研修コースも、評価システムを高度化する上で有効であろう。

このコースにおいても、ケーススタディ、事例のかたちで彼等の経験が紹介されることが望ましい。このような分野においても一般論が有効であるわけではなく、深い多様な経験に裏づけられ喚起される知識が有効であると考えられる。この場合、1つのシリーズとしては3日程度の研修期間を要する内容が想定される。

表 5-5 研究開発評価システム・コース

-
- ・ 研究開発評価システムの設計と運営
 - ・ 多様な評価システムのベンチマーク
 - ・ 研修期間：3日程度

[評価事例]

- ・ AAAS : 米国州政府の研究開発評価システムの設計と運営
- ・ テクノポリス : 欧州各国の研究開発評価システムの設計と運営
- ・ InDyne : NASAの科学振興プログラムにおける評価システムの設計と運営

・ 研究開発プログラムの運営（従属プロジェクトの評価）

研究開発に係るリサーチ・エージェンシー（省庁等）における研究開発評価の大部分は、研究開発プログラムの下で展開される従属的プロジェクトの評価であり、言い換えれば、これが評価の運営内容そのものである。米国の NSF、NIH、DARPA、NOA、NASA 等に在席するプログラム・マネジャーは、多様な機関内外のチェック・

アンド・バランス体制の下で緊張しつつ、プログラムの質を向上させるために、様々な工夫を凝らして、その見直し及び日常的な運営に携わっている。このような実務者の経験も貴重であり、彼等を講師とする研修コースもユニークなものとして構想できる。彼等の中でも特に評価の高いプログラム・マネジャーたちのリストを表 5-6 に掲

表 5-6 研修コース講師の候補者：
米国の行政内部のプログラム・マネジャーから

-
- | | |
|--|--|
| ▶ Leo Young
Consultant to the Research Office
Office of Secretary of Defense
Department of Defense | ▶ Dr. Gregory C. Tasse
NIST Program Office
National Institute of Standards
and Technology (NIST)
Department of Commerce |
| ▶ Jane A. "Xan" Alexander
Deputy Director
Defense Advanced Research
Project Agency (DARPA) | ▶ Paul N. Doremus
同上 |
| ▶ Dr. Ronald N. Kostoff
Office of Naval Research
Department of Defense | ▶ Stephanie Shipp
Director
Economic Assessment Office
NIST
Department of Commerce |
| ▶ William J. Valdez
Associate Director
Office of Planning & Analysis
Department of Energy | ▶ John Uzzell
Director
Division of Evaluation
Office of Science Policy
National Institute of Health (NIH) |
| ▶ Rich Beck
Director
Resource Analysis Division
NASA | |
-

げる。

米国だけではなく、ヨーロッパにおいても、例えば英国の Research Council あるいは Research Council を統括する DTI 内部の RC 統括部署等のプログラム・ディレクターは、この面での経験が豊富である。

表 5-7 研究開発プログラム・コース

- ・ 研究開発プログラムの運営
- ・ 研究開発プログラムのベンチマーク
- ・ 研修期間：3 日程度

[評価事例]

- ・ NSF, NIH, DARPA, NASA, NOA等 : 米国リサーチ・エージェンシーにおける研究開発プログラムの運営
- ・ DTI, EPSRC等 : 英国における研究開発プログラムの運用

また、それぞれのプログラムの運営それ自体の紹介にはそれほどの時間は要しないが、多様なプログラムの運営や評価ポリシーの違いを比較検討することは、自らが取り組まなくてはならない研究開発プログラムの運営を高度化する際に、非常に有益な情報を生み出すであろうことに留意しておきたい。このような経験豊かな実務者たちを講師とする研修コースをシリーズとして考えるとすれば、1シリーズ3日程度を予定するのが適切であろう。

以上述べた実務的な事例を対象とする研修コースの場合、行政内部やファンディング機関あるいは研究所の評価マネジメントに携わる職員等が研修コースの主要な対象者となる。さらに、シンクタンクの担当者にとってもこのようなコースは有益である。

取えて区分するならば、第一の社会経済性分析は、シンクタンクの研究者を中心とした受講生であるべきであるし、また、第三の研究開発プログラムの運営は、マネジメントポストにある者が中心であるべきであろう。また、第二の R&D 評価システムの設計と運営は、その両者にとって有益であると考えられる。

・機関評価の扱いについて

このような事例を中心とした研修コースの中に機関評価のコースは敢えて設定していない。その理由としては、研究機関の位置付けや内容は、機関ごとに極めて多種多様であり、それらを発展コースの事例として学ぶ意味と効果がどこまであるか疑問であるからである。

機関評価に関しては、むしろ本格的入門コースの中で十分にその取り扱いについての紹介がなされるべきであると考えられる。その前提として、研修コースの担当講師は、幾つもの機関評価に関する知識を有する者が望ましい。そのような比較された経験知が、自らの機関評価を行う際に有効な知識となると考えられる。別の観点から言い換えると、機関評価こそ、その研究機関に固有の観点からの評価システムが設計され運営されるべきものであるが故に、あまりにも個性的な事例を学ぶことは、かえってそれによって歪められるおそれが強い。このような場合には、相対化した知識を基盤として個別に設計に取り組む考え方がふさわしい。

B. 海外研究事例に学ぶシリーズ

欧米には、サイエンス・ポリシーや技術経営あるいはリサーチ・マネジメントといった、より広い専門性をカバーする一群の研究者に加えて、研究開発評価を専門とする、より専門的な研究者が既に存在し集積している。彼等のキャリアパスは多様である。比較的多いケースとしては、大学や研究所で経済系のバックグラウンドを持って育った研究者、さらには当初行政の実務者として経験を積んだ者が大学に席を移し、そこで研究経験をさらに積んだ者、一部には国の評価機関の中にあつて評価を専門とする業務に携わりつつ研究者としての質を深めたという者等がみられる。自然科学の研究者や工学の専門性を出発点とする研究者は、意外にそう多くはない。このことが示唆するように、研究開発評価の中心的課題は、科学技術の局面の評価にあるのではなく、社会経済性にあり、また、それらを含めたマネジメントに主題があると理解される。

このように専門性を一段と深めた研究者たちは、近年、研究者ネットワークを形成し、そのネットワークを現在も広げつつある。ヨーロッパにおいて研究者ネットワークの形成を促す契機となったのは、欧州委員会ECのフレームワーク・プログラムの評価を巡る議論からである。ECでは第4次フレームワーク・プログラムの評価において、初めて、評価専門家たちの助言を得ることが試みられた。当時は未だ評価専門

家はそれほどヨーロッパにおいても数多く存在していなかったが、その後各国にそのような専門性を備えた専門家が誕生し始め、そして欧州委員会が委託する評価プログラム、あるいはフレームワーク計画の評価委員会のサポートを舞台として彼等が活躍

表 5-8 海外研究事例に学ぶシリーズ

- ・ 最新の研究成果
- ・ 実務的事例の相対化

[研究紹介]

・ 欧米の研究者ネットワークメンバーによる

Stefan Kuhlmann(独) / Philip Shapira(米) /

Susan Cozzens(米) / Philippe Laredo(仏) /

Luke Georghiou(英) / Terttu Luukkonen(フィンランド) /

Anthony van Raan(オランダ)

し、相互の研鑽を深めるためのネットワークが形成されてきた。このような経過の中で、比較的初期の成果がまとめられた文献として、Laredo たちの編纂になる“The Strategic Management of Research and Technology”が挙げられる。これは、初期のテキストブックとしての機能も果たすことになった。

この本の中では、欧州共同体の専門家によるピア・レビューの現状紹介に始まり、評価に用いるサイエントメトリックス（科学計量学）の入門的解説があり、さらに戦略的マネジメントの解説があり、公共的研究開発の及ぼす経済的インパクトの評価に関するいくつかの論考が収録されている。そして、最後に欧州におけるサイエンス・プログラムのネットワーク効果の計測を論じ、技術－経済ネットワークと構造的効果の分析についての考察で終わっている。また、この本の中で論じられている、「(戦略的) マネジメント」と「評価」の区別をどうするかという論点は極めて深い意味を持っている。最終的には評価を論じるのみならず、plan（企画・立案）－do（実行）－see（評価）のサイクル全体についての理論構築をし、その中の see の部分に評価理論を位置付ける包括的視点による研究が必要となる。なぜならば、狭義の評価の局面で良い評価をしても、それが実際のプロジェクト運営にうまく反映されるのかされない

のかは全く別のイシューであり、評価をプロジェクトにうまく反映させる仕組みが構築されて初めて評価そのものの存在意義が出てくると言えるからである。また一方で、評価においてどこまで具体的に強制力を持つ提案を行い、企画・立案の領域に踏み込んで良いかと言う点も極めて微妙な問題である。

その後、第5次フレームワーク計画の評価システムを高度化するための議論にこれらの知見が引き継がれ、現在は第6次のフレームワーク計画の評価法を高度化するための議論へと発展してきている。それぞれの検討は同様の成果物としてまとめられている。

一方で、自国の研究開発評価機能を高めるために、欧州主要国、特にドイツにおいては、評価研究者と評価担当実務者等のコミュニケーションを深めるための国際会議が1997年以來持たれてきた。その会合が発展する形で、評価体制の整備が遅れている中欧諸国の評価実務者たちを招待して行う、第2次の国際的会合へとネットワークを広げてきた。さらにこの流れはヨーロッパ域内だけではなく、米国の研究者との交流を意図したネットワークに発展し、2001年に初めてそのような国際会議がドイツで持たれた。それ以後、米国とヨーロッパにおいて交互に欧米の評価専門家たちによる国際会議が持たれることとなり、それぞれの視点からの研究成果の報告・交流が行われている。

このようにして広がってきたネットワークの中心は、研究開発評価の研究者たちであるが、その周辺には科学技術政策や研究開発管理という、より広い専門性を有する専門家たちも配置され、広い見地から評価の問題と取り組む枠組みを与えることができていると言える。このような研究者ネットワークの発展過程と国際会議の成果物を表5-9にまとめた。

これらの活動の成果物自体が評価人材の重要な学習文献資料である。

また、研究開発評価の専門的研究者と目される研究者のリスト及びホームページ等を巻末の参考資料2にまとめた。また、彼等が属する何種類かのネットワークのホームページも同様に参考資料2にまとめた。

ここに記した評価の専門的研究者らは、いずれも教授級のシニアな研究者であるが、この他に次の世代を担う若い研究者たちも数多く存在している。もちろん彼等もこのようなネットワークの中に包摂されているが、ここでは具体的な氏名を挙げることはしていない。

このような海外研究者を講師とする研修コースを設定する場合に、注意すべきことは一般論に陥ることを避けるべきことである。むしろ、一般論を取扱うとしても、そ

表 5-9 研究者ネットワークと国際会議

1. European RTD Evaluation Network

欧州委員会の下で運営されている欧州域内の評価ネットワーク。各国1名のメンバーが決められていて、多くは行政官だが、研究者の場合もある。年1回域外にも公開された国際会議が開催される他、内部メンバーのための情報交換が行われる。ホームページはない。

2. The World Research Evaluation Network

<http://www.reseval.net>

ジョージア工科大学の Susan Cozzens がサイトの更新を受け持ち、評価に関係した国際会議、研究者リスト、リンク、国際会議のプロシーディングス、ワーキングペーパー等が掲載されている。ちなみにこのサイトはNSFがサポートしている。

3. International Evaluation Research Group (INTEVAL)

<http://www.c3e.fr/inteval>

1986年以來、欧州を中心とした20ヶ国以上の評価に関連した研究者や実務者が年1回会合を持ち、研究成果等の出版計画について相談をする。15年間の活動成果が書籍として集積されている。

4. Six Countries Programme – The International Innovation Network

<http://www.6cp.net>

1975年に設立されたイノベーション研究と政策形成に関わる専門家や政策担当者からなる国際的なネットワーク。カナダ以外は欧州各国の合計11ヶ国からなる。年2回ワークショップを開催し、成果はウェブサイト等に掲載される。主題が評価に限定されてはいないが、研究者ネットワークとしては重要。

5. EU-US 評価国際会議

2000年9月に欧州で開催された後、次年度、米国で開催された。欧米の評価研究者が一堂に会した。

- ・ Learning from S&T Policy Evaluation (September 11-14, 2000)
- ・ Research Assessment : What's Next? (May 21-23, 2001)

これはイントロダクトリーな限定された時間に限るべきであり、彼らに期待する本質的な課題は、評価に関する新しい整理された研究結果としての概念や知識を受講者に理解させることである。そのためには、研修内容として、実務者の個別経験だけでは到達することが困難であるような、実務的経験を相対化し論理化して捉え直された、より広い普遍的な知見であるとか、あるいは逆に、対象を絞りコンセプトを深めた、より深い本質的な知見こそが有益であろう。このような知見は、直ちには実務的レベルに反映させることが困難である場合が多いであろうが、しかし一方で、評価システムの本質的な認知の枠組みを意識して活動する上でも、そのような先導的な知見に接する機会もまた必要である。

欧米における研修コース

欧米においては、このような研究者が講師陣の中心になっている入門的な研修コースが、数はそれ程多くはないが大学等に付設され、展開されている。その概要を表5-10にまとめ、代表的な事例についての詳しい説明を巻末の参考資料1にあげた。

これらの研修コースの特色は、本報告書で位置付けた本格的入門コースとしてのマネジメント・コア・コースと異なり、研究開発マネジメントの枠組の中のひとコマに位置付けられていることであろう。つまり、評価についての専門的深さが十分あるとしても、その周辺知識が別の講義によってもたらされることを前提としている点に違いがある。別の言葉で言えば、そのコースを修得すれば、評価に関しての全ての関連知識が得られるというわけではなく、かなり限定的な評価に関わる知識が取扱われていることが多いと言える。

我が国の現状では、研究開発マネジメントなどに関わる評価の周辺知識を高等教育の中で取扱うコースが整備されていない。したがって、そのような周辺知識無しに、欧米で行われている評価研修コースを受講したとしても、必要とされる知識の全体像を修得することにはならないであろう。このように、研究開発評価の研修コースを設定する場合には、その研修の前提となる受講者の知識がどれほどの広さと深さがあるかということを十分見極めた上で、カリキュラム内容が検討されるべきである。そのような知識が十分に得られない海外の研究者による入門的コースのみに頼ることは、やはり危険であろう。

本報告書で構想した、海外の研究評価研究者による研修コースは、むしろ、実務者のより高度な知識を得るための研修コースと位置付け、発展コースをさらに発展させたものと考えべきである。これは、同時に、若い研究評価研究者のレベルアップに

表 5-10 欧州における研修コースの概要

▶ マンチェスター大学 PREST の教育コース内容(2002 年 1 月)

1 日目	イントロダクション(研究評価の概要) 評価の理論的基礎 ロジックチャート(講義) ロジックチャート(演習)
2 日目	ロジックチャート(演習)つづき ロジックチャート(講評) 経済的効果の評価
3 日目	ピアレビュー ビブリオメトリクス(講義) ビブリオメトリクス(演習、講評)
4 日目	ケーススタディ: Research Assessment Exercise 質問紙調査(講義、演習) ゲスト講師: 多レベルの評価(コンサルタント) 質問紙調査(講評) 社会的効果の評価
5 日目	評価のインパクトと利用 ゲスト講師: EU における研究開発プログラム評価 ゲスト講師: 英国 DTI における評価と政策形成

▶ トウエンテ大学の教育コース内容(2000 年 10 月)

1 日目	評価の社会的文脈	イントロダクション 評価の社会的文脈 コース概要
	評価のデザイン	評価のデザイン 演習 1: 目的からデザインへ(ロジックチャートの演習と講評)
	方法のカタログ	事例: オランダの大学評価におけるピアレビュー
2 日目	方法のカタログ (つづき)	方法のカタログ(概要) 中間的インパクトとしてのネットワーク 「社会的質」の評価 経済的インパクトの評価 事例: EU の研究評価 演習 2: 評価の方法の総合的利用(演習と講評)
3 日目	政策のための評価	評価データを用いた研究政策の改善 事例: 政策的介入の方法としての評価 因果関係の理解 演習 3: 評価レポートの作成
4 日目	政策のための評価 (つづき)	演習 3(講評) コースのレビュー 将来の研究評価

も役立つものとなるであろう。

このように、1人ないし2人の海外研究者が担当する研修コースとしては、1つのシリーズが3日程度あれば一応十分であると考えられる。

第6章 転換コース案

第5章で検討した研修コース案が評価関連部署ないし評価関連機関に在籍する者を対象とするのに対して、本章で扱う転換コース案は、研究開発部署等のそれ以外の部署に在籍するものが、評価の専門性を身につけ、評価関連部門に配属されることを前提とした研修コースである。「転換」という意味は、キャリアパスの転換を促すという意味である。

我が国における緊急の課題としては、既に設置され始めた研究開発評価関連部署ないし機関の在職者のスキルアップを図ることと同時に、なお圧倒的に不足している、それら実務的専門家を補うために、研究開発セクターから研究開発評価やそれに関連するマネジメントの専門部署へ人材を送り込むメカニズムを強化することである。米国の場合には、60年代末以降、特に72年以降に、そのような転換メカニズムが計画的に強化され、そのメカニズムにおいて養成された人材が、現在中央省庁や議会のスタッフとして活躍している。また、欧米の高等教育機関においては、大学での専門分野と、大学院での専門分野を大きく変えることは一般的であり、自然科学や工学の専門教育を受けた大学卒業生が、経済や経営、あるいは政策の大学院に進学し、キャリア転換を図るためのトレーニングを受けることも普通に見られる。米国のAAASが主催する転換コースは、そのような高等教育機関で行われる転換よりもさらに焦点を絞って、科学技術行政に人材を送り込むメカニズムを用意していると言える。

ここでは、マネジメント支援者養成コースとエキスパート養成コースの2つの転換コースを構想し提起した。このうち、マネジメント支援者養成コースをまず第一に整備すべきであろう。エキスパート養成コースが期待する人材養成メカニズムは、この第一のマネジメント支援者養成コースから派生する人材をも含むものであり、また、エキスパート養成コースそれ自体は、より多くはシンクタンクのためであることを考えるならば、シンクタンクとの協力のもとに展開することも考えられるからである。

6.1 マネジメント支援者養成コース

AAASの転換コース（参考資料1事例A-4）をモデルとして、我が国に適した転換コースを構想してみる。

表 6-1 AAAS の転換コース

- ▶ 第1日
 - Overview of Orientation and the Fellowship Year
 - The American Experiment in Government
 - Introduction of AAAS Fellowship Program Staff
 - Introduction of 2001-02 AAAS Science and Technology Policy Fellows
 - Welcome Reception

- ▶ 第2日
 - Where Does Science Fit in Public Policy? Domestic and International Perspectives
 - Differences in the Worldviews of Scientists and Policymakers, and Implications for Communication between Them
 - National Economic Policy
 - Concurrent Sessions: Ethical Requirements in Congress
Ethical Requirements in Executive Branch Agencies

- ▶ 第3日
 - Introduction to Federal Budget Procedure: Or, Why You'll Never Understand the Policy Process Unless You Understand the Budget
 - The Budget Resolution Process: The Politics of Budget Decision Making
 - Reflections on 53 Years of Science and Technology Policy

- ▶ 第5日
 - A Conversation with the President's Science Advisor's Staff
 - History of Science Advice to the President
 - A Lunch Conversation on the Office of the Presidency
 - Perspectives on the Congress
 - Legislative Process

- ▶ 第5日
 - **Legislative Track**
 - Overview of the CRS
 - Science, Technology and Congress: Overview and Selected Issues
 - Lunch with CRS Staff
 - Understanding the Congress
 - Conclusion & Wrap-up

 - **International Track**
 - International Science Activities in Federal Agencies: Case Studies

- The International Development Cooperation System
- Developing Countries: Why Should We Care?
- Science in Foreign Policy: A Case Study
- **Executive Branch Track**
- The Role of Science Fellows in Agencies Filled with Scientists
- Coordination of Domestic Science in Federal Agencies
- the Role of Policy Analysis and Information in the Executive Branch: Case Histories
- The Role of Regulatory Agencies in the Executive Branch

- ▶ **第6日**
- Foreign Affairs: Current Issues and the State Department’s Role
- Development Assistance: Current Issues and USAID’s Role
- The Nature of the Federal Bureaucracy: Its Structure, Function, and Culture
- Between Traditions and Modern Challenges: Lithuania’s Natural Environment in the 21st Century

- ▶ **第7日**
- **International Track**
- The Role of NGOs
- U.S. Leadership Role in Global Affairs: Use It or Abuse It
- Lunch with Member of Congress
- U.S. Foreign Policy: An Interactive Workshop

- **Legislative Track**
- Perspectives on the 107th Congress
- How a Congressional Office Works
- Lunch with Member of Congress
- U.S. Foreign Policy: An Interactive Workshop

- **Executive Branch Track**
- Federal Advisory Committees
- Government Performance and Results Act
- Lunch with Member of Congress
- U.S. Foreign Policy: An Interactive Workshop

- ▶ **第8日** – Lobbying in Washington
 - The Energy-Climate Challenge: and What to Do about It
 - Science in the Judicial Branch

 - ▶ **第9日** – Science and Technology at the Department of Defense
 - Overview of the National Academy of Sciences
 - How to Operate in a Bureaucracy with Intelligence and Integrity

 - ▶ **第10日** – Briefings on Current Issues from a Disciplinary Perspective
 - Washington’s Think Tank
 - When Politics and Science Have Different Agendas: The FDA and Tobacco Regulation
 - Reporting Science and Science Policy News in Washington

 - ▶ **第11日** • **Congressional Fellows**
 - Preparing for Placement
 - Legislative Resources

 - **USAID Diplomacy Fellows**
 - Overview of USAID
 - The Role of AAAS Fellows at USAID
 - Information Regarding Stipends, Taxes, Travel and Travel Vouchers

 - **Environmental Fellows**

 - **USDA/FSIS Risk Policy Fellows**

 - **All Other Fellows**

 - ▶ **第12日** • **Congressional Fellows**
 - Meet at Placement Office

 - **USAID Diplomacy Fellows**
 - USAID Budget Process
 - Strategic Objectives and R4’s
-

従来、我が国でも民間企業から行政機関へ人材が派遣されることは一般に行われてきた。科学技術関連部門には、民間企業の自然科学や工学をバックグラウンドとする人材が派遣されていた。このような人材と AAAS が行政機関や議会に送り込むスタッフとの違いは次の 2 点にある。第一は、送り込まれた人材の主体性に関わることであり、我が国の従来の派遣メカニズムでは、行政機関の補助者としての能力が彼らに期待されていたのに対して、米国の場合には、補助者ではなく、科学技術の専門性を備えている有力な支援者と位置付けられていることにある。また、第二には、従来の我が国のシステムでは、派遣される際に業務に関わるなんらの研修も行われていないのが実情であるが、米国の場合には 2 週間の非常にインテンシブな研修が行われていることである。

また、さらに付け加えていえば、AAAS の場合には、派遣元はリサーチ・コミュニティであり、学会の会員からの応募に基づき学会内部で選考された転出希望者である。具体的には大学や国立研究機関の、一般的には博士号を所持する研究者が大部分であるのに対して、我が国の場合には大学からの派遣は、特殊な例を除いて一般的ではなく、多くは民間企業からの派遣であり、本人の意思に基づく人事ではなく企業の人事方針に基づくものである。また、民間企業からの派遣者が博士号を所持していることは必ずしも一般的ではない。

最近、大学からの派遣、大学から行政機関への一時的な移行がより広く行われるようになったが、未だその数は多くはない。

新たに転換コースを我が国において設定するとすれば、次のようなものとなるべきであろう（表 6-2）。第一は、従来行われていた出向メカニズムをベースとする転換コースであり、出向者の自覚とインセンティブを高めるために、公募システムと選抜制度を関門として選抜した後に、集中的な研修コースを経て配属されていくメカニズムである。いわば、派遣元の論理から派遣者の論理に置き換える必要があると考える。第二のメカニズムとしては「転換」であり、従来の職場を離れて、期間を限定した雇用形態で行政内部に中途採用されていくキャリアパスである。この場合にも、やはりインテンシブな研修コースが必要である。

第一の出向メカニズムによる場合の対象者としては、例えばプログラム・ディレクターの適任者のようなシニアな研究者から、助教授クラスの中堅の研究者まで、幅広い層が受け入れられるべきであろう。出向メカニズムである以上、一定期間、例えば 2 年間の後に出向元に帰還することになるが、改めて第二のメカニズムにより、職を辞して中途採用されることも継続的なプロセスとして許されるべきであろう。また第

二の期間限定雇用の転換者が、期間を終えた後の職場として、他の省庁や米国におけるように、議会スタッフ、さらにはシンクタンク等を含めた、より広い受け皿へと専門性を活かして職を得ていくメカニズムが考えられる。このような流動的なメカニズムの中で広い経験を蓄積することが、この分野の専門家としての技能を深めるためにも望まれることである。米国の場合には、その転換者のほぼ三分の一がワシントン近辺に集積している。また、三分の一は派遣元に戻るが、派遣元をベースに、多様な人的ネットワークを活かし、より広範な研究分野での活動に移るといった状況が見られる。追跡調査によると残り三分の一は、このような転換が継続しなかった人たちである。例えば、純粹の研究者に再び戻るといったようなケースがそれに含まれる。

表 6-2 転換コース案の概要

1. 出向コース

- ・ 出向元の組織内で公募し選抜
- ・ 集中転換研修(2週間程度)

[研修内容]

- －行政組織の概要
- －政策形成過程
- －予算策定プロセス
- －予算執行プロセス
- －行政の主要課題
- －派遣先の特殊問題(パラレルセッション)

2. 転換コース

- ・ 任期付ポストに対して公募
- ・ 集中転換研修(2週間程度)
 - 上記の他にキャリアアップメカニズムの紹介等「雇用」関係により生じる問題の説明
- ・ 年間数回、全省庁的にまとめて行う

このような転換コースのカリキュラム内容としては、行政メカニズムのイントロダクトリーな説明に始まり、マネジメントの主要な枠組み、そして、戦略形成や評価といった特定の課題に関する、より深い研修や演習が想定される。このような内容をもつ研修期間としては2週間ないし3週間程度を要するであろう。なお、米国の場合には、終身雇用の枠組が殆ど実在しないために、出向制度がそれ程普及していない。AAASの場合には、それに代わる制度として、トレイニー（研修生）という形態を取り、その研修生に対しては、AAASが中間組織として資金を他から得た後、トレイニーの給与を補填するメカニズムが採用されている。つまり、研修生を受け入れる政府機関がグラントを整備し、AAASにそのグラントを出した後、AAASがそのグラントを基にして、研修生の給与としてそれを支出するメカニズムが採られている。また選考過程としては、米国の場合には、サイエンス・コミュニティからの人材派遣が主流であるため、有力学会が学会ごとに研修候補生を選抜し、その選抜された者が、AAASの研修コースに集められて研修を受ける形を取っている。この研修内容と研修体験記については、巻末の参考資料1事例A-4を参照されたい。

6. 2 エキスパート養成コース

社会経済性分析を典型例とするような分野の専門家を養成するコースとして、エキスパート養成コースが設定されるべきである。このコースの受講者は、研究開発に関する評価を対象に含むことから、大学や研究機関において自然科学や工学等をバックグラウンドとした研究者であるべきであろうが、その転換先として国の政策経営調査研究機関や民間のシンクタンク等へ配属されるような、研究者や技術者の転換研修コースと位置付けられる。そのメカニズムも出向制度をベースとする人材と、いわゆる転職・転換をベースとする人材に分けることが出来る。このコースの場合には、行政機構への一般的な導入よりも、社会経済性分析等の専門的な学問分野の行政機構への紹介が中心となるものであるべきであろう。また、先程とは逆に、このエキスパート養成コースをひとつのリザーバーとして、そのリザーバーの中から、マネジメント支援者養成コースへ進み、行政内部で活躍するメカニズムも考えられる。

このような転換コースにおいても2~3週間の研修が最低必要であろう。

第7章 高等教育専門課程

7. 1 科学技術経営政策研究課程（大学院修士課程）

科学技術の経営や政策を主題とする高等教育課程としては、実質的に該当するのは大学院課程となる。米国の場合には20程度のコースが実在しており、その実績についてはAAASが期間をおいて追跡調査をしている。特殊な事例を除くと、主に60年代後半から70年代初期にコースが開設されており、当初の卒業生たちは既に定年・引退の時期を迎えたという歴史をもっている。ヨーロッパ各国でも、この種の専門的なコースは早くから開設されており、その卒業生により、行政機関やシンクタンク等の外部支援機関の専門性が維持されてきているのが実情である。

我が国の場合には、このようなコースは、まだ端緒的分散的整備に留まっており、殆ど開設されていない状況にある。ただし、大学でのこの面の整備も重要であるが、現在の評価人材をめぐる火急の課題に対応することからいえば、先に述べた研修コースや転換コースがまず必要である。大学の整備は直ちに手をつけるとしても、その卒業生が実務にまで入るためには数年を要し、時間的な観点だけからみても、このメカニズムのみに依存するわけにはいかない。

海外における代表的なこの種の大学院コースのリストを表7-1にまとめた。このリストは、技術経営のみを専門とするコースは、ほぼ除外しており、何らかの形で科学技術政策を含むコースを主として取り上げた。技術経営のみのコースは、この数倍から一桁程度多く存在している。

7. 2 大学付設の短期研修コース

大学に、第5章や6章で述べた研修コースや転換コースが付設されることもある。欧州におけるその具体的事例について既に第5章で述べた（表5-10）。ここでは表7-2に、我が国で展開すべき大学付設の短期研修コース等のあり方についてまとめる。

この種の研修コースの特徴としては、大学に付設することにより大学の正規のコースとの互換性を持たせた設定形態が可能なことであろう。つまり、単位取得を可能とし、正規コースの学期制度に合わせて半年ないし一年の研修コースとして設定することも考えられる。その場合には短期ではないが、例えば土曜日あるいは休日のみに講

表 7-1 科学技術政策関連大学院コース(米国)

▶ **Programs in Science and Technology Studies: 科学技術論**

- Cornell University
- Rensselaer Polytechnic Institute
- Virginia Tech

▶ **Programs in Science and Technology Policy: 科学技術政策**

- The George Washington University
- University of Minnesota
- Princeton

▶ **Public Policy Programs: 公共政策の一部門として**

- George Mason University
- Georgia Institute of Technology
- Rutgers

▶ **Engineering/Public Policy Programs: 工学と公共政策**

- Carnegie Mellon University- H. John Heinz III School of Public Policy and Management
- Carnegie Mellon University - Engineering and Public Policy
- Massachusetts Institute of Technology

▶ **Other/Interdisciplinary Programs: その他の学術的プログラム**

- Arizona State University (Law, Science, and Technology; JD)
- University of Delaware (Energy and Environmental Policy; Ph.D. in Technology, Environment and Society)
- Georgia Institute of Technology (History of Technology, MS and Ph.D.)
- University of Oklahoma (Science and Public Policy-joint degree)
- University of Washington (Public Administration and a number of joint programs)
- Washington University (several joint programs)

AAAS のホームページに、受験生向けの「科学技術政策関連大学院コース」の案内コーナーがあり、具体的な詳しい情報にアクセスできるようになっている。

(<http://www.aaas.org/spp/sepp/sepsl.htm>)

義のコマ配置を行い、社会人を受け入れるコースとして位置付けることも考えられる（パートタイム研修コース）。また、一般勤務者のキャリアアップのためのコースと位置付けることもでき、他の隣接科目に対する同様のメカニズムによる単位修得等を合わせて、新しい領域の専門家へ転換していくキャリア転換のきっかけとして位置付けることも出来るであろう。第5章や6章の研修・転換コースに比べると、ここで例示したように多少期間を伸ばしたメカニズムを導入することによって、より本格的に転換を図る契機として、この種のコースを利用することが可能となる。それは、大学の教育環境を生かすことによって始めて可能となる。

表 7-2 大学付設の短期研修コースの設置形態案

1. 短期集中研修コース

- ・ 大学の休暇期に2, 3週間程度設定
- ・ 単位等の資格取得とは無関係

2. パートタイム研修コース

- ・ 学期に合わせて運用
 - ・ 休日ないし夜間に開講
 - ・ 単位制に適合させ資格取得も可能とする
-

第8章 評価人材整備のための施策

1. 多様な研究開発評価人材に見合った養成メカニズムが必要

研究開発評価の多様な局面を担うためには、専門化した多様な評価人材が必要である。関係する人数の多さの順にそれらを列挙すると、「科学技術ピアレビューア」、「一般の評価マネジメント人材」、「社会経済性ピアレビューア」、「評価マネジメント・コア人材」、「評価エキスパート」等である。したがって、それぞれの評価人材に見合った多様な研修コースを目的に合わせて設定する必要がある。表 8-1 に、これらの評価人材がそれぞれ履修すべき研修コース名を示した。

表 8-1 研究開発評価人材とスキルアップのために必要な研修コース

主な所属機関	評価人材の種類	マネジメント・コア・コース (本格的入門コース)	エキスパート・コース (発展コース)	
			海外実践 事例に学ぶ	海外研究 事例に学ぶ
行政機関・資源配分機関等	評価担当部署 (評価コア人材)	○	○	○
	評価実施部署 (評価一般人材)	○		
外部支援機関 (シンクタンク等)	エキスパート	○	○	○
研究機関	評価担当部署 (評価コア人材)	○	○	
	評価コアへの 転換人材	○	○	○
	行政機関等への 転換人材*	○	○	○
大学等	評価研究者		○	○

*この他に行政入門研修が必要

2. 先ず、「評価マネジメント人材」と「社会経済性評価人材」の養成から着手すべき

我が国で展開されつつある研究開発評価の状況と、研究開発評価人材の充足状況とを見比べ、さらに評価人材が担う職務内容の重要性を考慮すると、以下の順序で重点的に研修コースを設定すべきである。

(1) 「評価マネジメント人材」

評価マネジメント人材は、評価の要として、評価システムの設計と運用を担うべき立場にあるが、行政機関、資源配分機関、研究開発機関を問わず、現在この立場にある者の多くは妥当な研修を受講していない。近年全面的に実施されるようになったプロジェクトの事前評価や機関評価における混乱の回避は喫緊の課題であり、「評価マネジメント・コア人材」のみならず、「一般の評価マネジメント人材」にも対象を広げ精力的に研修機会を設定すべきである。

(2) 「社会経済性評価人材」

研究開発評価を実施するうえで、科学技術の質的評価に比し社会性や経済性の効果の評価することの方がはるかに困難である。それにもかかわらず、社会経済性評価のエキスパートはおろか妥当なピアレビューアを見出すことも困難である。従来から行われてきた研究開発者による社会経済性の評価は、楽観側に偏る傾向が著しく、有効性を欠いていた。我が国の評価体制におけるこの欠陥は一刻も早く是正されるべきである。

(3) 研究開発者からのキャリア転換

上記(1)、(2)で述べた専門家への研修者は、対象機関の在籍者だけでは、能力的にもまた意思的な動機の面からも、十分な適任者を得ることが困難であろう。特に行政機関や資源配分機関においては、研究開発の本格的な経験を有する者の中からこの種の適任者を見出すことはほとんど不可能である。このような理由から、研究開発者の評価専門家への転換を促す「転換研修コース」の設定も重要である。

(4) 「科学技術ピアレビューア」

科学技術の側のピアレビューアは、従来からまがりなりにも整備され、現在マスとして存在している。科学技術ピアレビューアの現在の本質的課題は、彼および彼女等の質的向上にある。しかし、評価に携わる研究開発者に対し、評価のための研修を本格的に義務付けることは現実的ではない。本報告書では、評価にたずさわる研究開発者が本来的な科学技術ピアレビューアの立場に立ち戻り、その専門とする科学技術領域に関してのみの価値判断に携わることを推奨している。このように整理した場合、

評価に携わる研究開発者に対して必要なことは、評価実務に関する簡単なガイダンスであり、それはピアパネルを実施するに当たって評価マネージャーによって説明されるべきものである。従って、本報告書では科学技術ピアレビューに対する研修コースを特には設定していない。状況によって必要になる場合には、(3)で用意する研修コースの一部を適用すれば十分であろう

3. 研究開発評価に限定しない広い能力開発であるべき

研究開発評価の実務は、ほとんど全ての場合評価の「応用問題」に対処することに等しい。評価の目的や状況に合わせて、評価の在りかたに対し創造的に取り組む必要がある。評価マニュアルの機械的な適用は厳に慎むべきである。この観点から、修得すべき能力とそのための研修内容として以下の各項を想定している。

(1) 評価のための実務的スキル

まず研究開発評価の中核をなすワンセットの実務的方法論を修得する必要がある。その内容としては、評価の推進体制の在りかたと評価の諸局面に関する全体的な知識の導入、評価の実務にとって重要な具体的な方法論の体系的な理解、評価対象のカテゴリー毎にまとめた評価体制の設計と運用の在りかた、そしてこれら全ての要素を包含する多様な事例に基づく演習等である。

(2) 手法を使いこなすための基盤的能力

(1)で述べた研修内容は、研究開発の評価に焦点が絞られている。これらの知識が研修内容の中核をなすことは言うまでも無い。しかし、このような知識を実務において使いこなしていくには、これらの知識の背景や基盤をなすより一般的であったりより深い知識体系について理解を進めておく必要がある。まず、評価の一般論である評価論、そして方法論の基盤をなすシステム論やシステム思考・論理思考等、さらには評価作業や評価対象が組織を基盤にしていることから組織現象を取り扱うための組織論等についての理解が必要である。

(3) マネジメント全体を見通す広い視野

研究開発評価は研究開発マネジメントの一部をなすが、評価のみをマネジメントから切り離し運用することはほとんど不可能でありまた無意味でもある。たとえば、事前評価では評価対象の上位にある計画やさらにその基となる戦略との関係が議論される必要があり、また途上評価や直後評価では評価対象の目標や目的が明確になっている必要がある。戦略、計画、目的、目標等の設定は評価行為の範疇にはなくより広

いマネジメントの枠組みの下での思考や運用のカテゴリーに属している。このような観点から、修得すべき知識を評価に限定するのではなく、研究開発のマネジメント全体を見通せる知識が必要であり、科学技術のマネジメント論を構成する一連の知識体系の理解が重要である。

4. 研修内容に相応しい仕組みと講師の選定を

研修内容は我が国の現在の状況を踏まえたものである必要がある。現在の我が国の評価状況は、歴史的経緯から、極めて個性的でありまた反面バランスを欠いたものでもある。これは海外の研究者や実務者の常識的理解の枠組みを超えている。そのため、研修内容のポイントと講師の選定に関しては、その状況に適合しまた相応しくするための特段の配慮が必要である。一方で我が国に集積されている評価事例や講師適任者の数を考慮すると、海外の講師に依存せざるを得ない面もあり、これらの事情を勘案し、各コースの内容と講師陣の在りかたに関し表 8-2 に示したように、入門的側面は我が国独自に対応し、また高度な発展的内容については海外講師陣に依存するという基本方針が妥当であると考えらる。

表 8-2 主な研修内容と妥当な講師像

研修コースの区分	主な研修内容	妥当な講師像
本格的な入門コース	<ul style="list-style-type: none"> ・評価の推進体制全般に関する理解 ・評価の諸局面の把握 ・評価の方法論の修得 ・評価推進体制の設計と運用 ・ケース・スタディと演習 	国内の評価事情に詳しい評価研究者や実務者
発展コース 海外実践事例に学ぶコース 海外研究事例に学ぶコース	<ul style="list-style-type: none"> ・独立大型プロジェクトの事前評価事例 ・プロジェクトの途上、追跡評価事例 ・研究開発評価システムの設計と運用事例の比較 ・従属型研究開発プロジェクトの事前評価事例の比較 ・最新の研究成果 ・実務的評価事例の相対化 	海外の評価実践経験の豊富な研究開発評価エキスパートおよび評価実践経験の豊富な研究者 海外の第一線で活躍している研究開発評価研究者
転換コース	以上の研究開発評価の研修内容の他に、行政メカニズムの理解のための研修内容を付加する	行政メカニズムの理解のためには、シニア行政官と政策研究者
大学付設のコース 短期集中研修コース パートタイム研修コース	基本的には、上記の「本格的入門コース」と「発展コース」の内容と同じ 上記「短期集中研修コース」の他に、その内容を使いこなし発展させるための「基盤的能力」の開発とマネジメント全体を見通す広い視野を獲得するための「経営学」関連学問	基本的には上記の「本格的入門コース」と「発展コース」の講師と同じ 上記の「短期集中研修コース」の講師の他に、システム論や経営学の専門家



参考資料 1. 事例

事例 A 評価人材養成プログラムに関する事例

- A-1 マンチェスター大学 PREST における教育コース
- A-2 オランダ・トゥエンテ大学における研究評価の教育コース
- A-3 米国ジョージア工科大学 (GIT) における教育コース
- A-4 AAA S 科学技術フェローのオリエンテーションプログラム

事例 A-1. マンチェスター大学 PREST における教育コース¹

A-1.1 教育コースの概要

英国マンチェスター大学の PREST (Policy Research in Engineering, Science and Technology : 工学科学技術政策研究所) では 5 日間の集中講義の形式で研究評価の教育コースを開講している。このコースはそもそもは PREST の 1 年間の修士課程(MSc) 中の授業の一つとして位置づけられているものであり、それを外部にも公開したものとなっている。とは言っても、この修士課程には各国の科学技術行政を担当する省庁から留学生が多く派遣されており、また、英国以外の国の省庁や研究施設からも実際にコース受講者が来ている。そのため、研究評価の実務を行っている現場でも役立ちうるように、実習を多く含めた内容となっている。

教育コースは毎年 1 月の半ばの週の月曜日から金曜日までとなっており、今年(2002 年)は 1 月 14 日～18 日であった。ほとんどの日は朝 10:00 から夕方 17:00 頃というスケジュールになっている。なお、PREST の学生以外の受講料については、EU 圏外の受講者は 800 ポンドであり(2002 年の場合)、EU 圏内であれば割引があるという。

教育コースの講師は、研究評価研究の第一人者である Luke Georghiou 教授をはじめとして主に PREST 所属の教員で構成されている。これは後述するオランダのトゥエンテ大学が各国の研究評価研究者を集めて行っているのとは異なる。その理由としては、このコースが PREST 内部の授業でもあること、PREST が研究評価研究を主要テーマの一つとする研究機関であるために評価研究・調査の経験や知識を有する教職員が揃っ

¹ PREST の教育コースについては、2001 年 11 月に Georghiou 教授および Dr. Boden に対してヒアリング調査を行った。また、科学研究費補助金「国際的通用力を持つ大学評価システムの形成に関する日、欧、米の国際共同研究」の一環として、2002 年 1 月に開かれた教育コースに実際に参加して調査を行った。

表1 マンチェスター大学 PREST の教育コース内容 (2002年1月)

1日目	イントロダクション (研究評価の概要) 評価の理論的基礎 ロジックチャート (講義) ロジックチャート (演習)
2日目	ロジックチャート (演習) つづき ロジックチャート (講評) 経済的効果の評価
3日目	ピアレビュー ビブリオメトリクス (講義) ビブリオメトリクス (演習、講評)
4日目	ケーススタディ: Research Assessment Exercise 質問紙調査 (講義、演習) ゲスト講師: 多レベルの評価 (コンサルタント) 質問紙調査 (講評) 社会的効果の評価
5日目	評価のインパクトと利用 ゲスト講師: EUにおける研究開発プログラム評価 ゲスト講師: 英国 DTI における評価と政策形成

ていることが挙げられる。ただし、実際の評価事例の紹介については、外部からゲスト講師を招いている。

また、このコースで対象としている受講者は、上述のように2種類の受講者が実際に受講していることから分かるように、行政組織や資金配分機関あるいは研究所などで評価に当たる実務者、および、研究評価研究や科学技術政策研究の分野のアカデミックな研究者や民間調査機関の調査研究者・コンサルタントなどを旨とする者である。そのため、教えられる内容はこれら2者に共通的に必要な基盤的内容となっている。

A-1.2 教育コースの内容

PREST の教育コースの内容は表1のようになっている。大きくは、次の内容に分けることができ、これら内容にゲスト講師による事例紹介が加えられて全体が構成されている。

- ・ 評価が必要となった社会的背景
- ・ 評価の理論的基礎
- ・ ロジックチャート
- ・ 評価に用いる手法
- ・ 評価の意思決定への利用

(1) 社会的背景

コース初日には、まずコースのイントロダクションの意味を含めて、研究評価が必要となってきた背景についての概要説明が行われる。すなわち、1980年代以降に英国や欧州全体で研究開発へ投資される公的資金が上限を見るようになり、各研究プロジェクトやプログラムの効果や公的投資の正当性を分析し、優先順位設定をすることが必要視されるようになってきた。この流れはその後の行政改革におけるニュー・パブリック・マネジメントの流れで強化されてきたものである。このように研究評価が制度化されるようになった背景をはじめに説明することで、「何のための研究評価であるのか」といったコースを学ぶ上での動機付けを受講者に行う。

(2) 理論的基礎

次に評価の理論的基礎が説明される。ここでは、研究開発の評価を行うシステム構成が説明される。これは、評価という営み一式がどのようなもので構成されるのかを簡単に説明するとともに、逆に、具体的に評価を行っていく上でどのような要素について考慮する必要があるかを考えるための基礎となる概念化であると言える。

ここでは、Gibbons and Georghiou (1987)などを参照としながら、評価のスコープとして研究の種類（基礎・応用・開発や戦略研究などの区分）、評価対象のレベル（政策・プログラム・プロジェクト）、評価のタイミング（事前・中間・事後・追跡）などをまず認識する必要があることを説明する。また、一般的な評価基準（効率性、効果・インパクト、適切性など）や、評価に関わる人材の違い（当該研究分野やそのアプリケーション領域の専門家、組織内外の方法論的専門家）が簡単に説明される。

研究開発の成果やインパクトを評価する際には、前提的概念として、特定の研究活動から得られる効果は単一のものではなく多様な種類が生じることを認識することは重要であるが、それらについても説明される。また、特定の種類の効果の分析手法の説明に入る前に、評価に共通して必要となる一般的手法が説明される。それらはデータ収集の方法（インタビュー、サーベイ、参与観察、統計など）やデータ分析の方法（ケーススタディ、モデリング、データ分析、指標、コストベネフィット分析）、評価のフレームワーク（前後比較（Becher 1997）、コントロール・グループ・アプローチ（Meyer-Krahmer 1997）、反事実条件比較、ロジカルフレームなど）である。

このように評価システムの構成要素がこの講義では概説されるが、ここでいう評価システムは若干狭い意味に限られている。評価システムは本来はそれを内包する意思決定システムとの関係で形成されるものであるが、そのような広い視点からいかに評価システムを構成するかについての説明は行っていない。これは、受講者として想定されてい

る者が、特定のプログラムやプロジェクトの評価を担当している（あるいは将来担当する）者であり、評価システムを意思決定の中にかに位置づけるかといった内容はそれら受講者の権限を越えたものになりうるためであるとも考えられる。ただし、最終日には評価結果を意思決定へと連結させるための評価のあり方についての説明はなされる。

また、講義ではより一般の政策評価やプログラム評価における評価理論の先行研究の紹介もされる。「評価」という行為はその定義自体も多様であり、それ故に評価が基礎を置くパラダイムも多様である。例えば、存在論的には事実は一つしかないのか、それとも事実は取得可能な情報を基に構築されたものでしかないのかが問題となる。同様に、認識論的には評価者は被評価対象から距離をおいて観察できるのか、得られる内容は調査のプロセスの産物でしかないのか、方法論的には調査により何かを明らかにできるのか、評価者が解釈をするだけであるのかといったことが問題となる。このような「評価哲学」の講義があるのは、受講者にはアカデミックな研究者を目指す学生も入っていることが影響していると思われる。また、同時に、近年の政策評価の導入によって主流となりつつあるアウトカム・アプローチと、より実施プロセスへの焦点を強く持つプロセス・アプローチの違いなども説明される。これらの講義は特定のパラダイムを推奨するものではなく、評価という営み自体が、「評価対象についての客観的な情報を得て、その成果の価値を判断する」といった単純な行為には縮減できないことを認識すべきであることを説明するものであると考えられる。

なお、この講義の中でも先行研究を読んでまとめるという形の実習は行われる。

（3）ロジックチャート

次に、上述した評価の理論的基礎と、次節で述べる特定種類の効果の測定のための手法との間をつなぐレベルの方法として「ロジックチャート」（あるいは「ロジックダイアグラム」）が説明される。これは PREST により開発されたものであるが、オランダのトゥエンテ大学の教育コースでも講義されているものである。また、PREST が 1999 年に受託した通産省の大プロのプロジェクト評価でも用いているために、日本でも知られている手法である。

「ロジックチャート」は、評価対象のプロジェクトやプログラムについて、階層構造上に上位レベルの目的、下位レベルの目的（目標）、活動、期待されるアウトプット、期待されるアウトカムという順で並べていき、それらの間の関係をリンクとして示す方法である。プロジェクトやプログラムの論理構造をこのように階層的に示すことは一般の政策評価でも提案されている。しかし、それらでは目的と期待されるアウトカムは同一のものとして区分されないで概念化されることも多いのに対し、PREST のロジックチャートでは当初目的から外れて生じる負の結果や、明示的に示されない効果（研究者

のスキル向上など)、スピルオーバー的な効果(社会経済的効果や他の学問分野への波及効果など)についても示しうるようにしているところに特徴があると言える。

ロジックチャートは評価対象の論理構造を明確にするための手法である。評価対象がプロジェクトであれプログラムであれ、それら対象には、より上位のレベルから導き出される目的があり、さらにそれを達成するための具体的目標が設定されているはずである。ロジックチャートにより、それら目的構造と実際の実施行為および成果との間の論理的整合性を明確にすることで初めて、評価の基準を設定することが可能となるのである。また、実際の行為過程はプロセス、アウトプット、アウトカムというようなフェーズがあり、評価ではこれらをそれぞれに評価することも求められる。そのため、このようにプロセスや期待される成果を明確にしておくことが、その後の成果測定の手法の選択にも必要となると言える。

講義ではこのロジックチャートについての概要説明を行った後に、EUによる国際共同研究プログラムである EUREKA を事例としてロジックチャート作成の実習が行われる。

(4) 評価の手法

次に評価に用いる手法が説明される。評価で用いる手法には前述したように、対象に関する情報や関係者の主観的判断を収集する基礎的な方法と、特定の種類の成果の測定を行う技法とに分けることができる。教育コースではこれらの双方が教えられている。

基礎的な方法としては、ピアレビュー、質問調査がそれぞれ時間枠をとって説明される。ピアレビューの講義では、その種類自体が様々にあることがまず説明される。例えば、特定分野の専門家がパネル形式で行う「直接的ピアレビューや」、それ以外のアプリケーション分野や産業界などの人間を含む「改善されたピアレビュー」、論文引用や賞の受賞などを用いる「間接的ピアレビュー」というように分けることができる(Gibbons and Georghiou 1987)。また、ピアレビューの一般的問題点(保守的傾向、学際分野へのバイアス、女性・新参者・競争者へのバイアス、判定のインフレ、分野間の比較の困難性、社会経済効果の判定の困難性など)などが解説される。また、質問調査の説明では社会調査法の授業で教えられるような、質問肢の明確な作り方の注意事項が説明される。ここでも、いくつかの例題に対して質問肢を作成する実習が行われる。

次に各種の効果を同定する方法として、科学技術的効果、経済的効果、社会的効果それぞれの分析手法の説明が行われる。

科学技術的効果については上述のピアレビューの説明とは別にビブリオメトリクス(計量書誌学)の簡単な説明と問題点の紹介がはじめになされる。受講者の PREST の学生の中には Science Citation Index などのデータベースを実際にさわったこともない

学生もいるため、次に、コンピュータールームにおいてそれらを用いた簡単な実習が行われる。

経済的効果の分析手法については、Mansfield(1980,1991)に代表されるような民間企業等における研究開発への投資効果分析にも若干ふれるものの、そのような金銭的な形で経済効果を分析することは困難であることを説明する。その理由の一つには、得られた特定の経済効果に対してその原因として挙げられる要素は当該研究プロジェクト以外にも複数あることが通常であるし、逆に一つの研究プロジェクトの影響も一つの経済効果としては測定できないことがあげられる。そのため、実際にどのような効果が何を原因として生じているかを各評価対象に即して分析することが求められるのであり、単純にインプットとアウトプットを限定的に測定して分析するだけでは意思決定に十分利用しうる信頼性ある評価結果を行うことはできないためである。

このような理由から、経済効果分析の授業においては、まず経済効果の種類（製品やプロセス、雇用促進など）を紹介した上で、「スピルオーバー」概念についての説明が行われる。スピルオーバーとは、直接当該プロジェクトやプログラムに参加していないアクターへの効果であり、その効果が大きいことこそが公的資金を用いた政策介入の正当性につながる。スピルオーバーには知識のスピルオーバーや、消費者余剰に見られるようなマーケットスピルオーバー、ネットワークスピルオーバーなどが挙げられる。

さらにスピルオーバー概念と同時に政策介入を正当化するために、アディショナリティ概念(Lukkonen 1998, 2000)も重点をおいて説明される。特に民間企業への公的投資について、その公的投資が民間企業独自の投資行動へ与える影響や行為変化の誘因を分析することにより政策介入の効果を分析するためにアディショナリティ概念が必要となる。

これらの概念を紹介した上で、経済効果を分析する一般的方法としては、「改善されたピアレビューmodified peer-review」あるいはメリットレビューと呼ばれるようなピアレビュー方式や、レトロスペクティブな跡付け、指標利用、生産関数、消費者・生産者余剰などがあることが簡単に説明される。また、特に近年は new economics of science (Dasgupta and David 1994) と称されるような科学社会学的視点の導入により、Techno-economic Network (Callon et al.1997) や knowledge value collectives (Bozeman 1999) に見られるようなリンケージの評価に焦点が当てられることも紹介される。

また、フランスのストラスブール大学が開発した BETA 法(Bach et al 1997)については詳細な説明がなされる。BETA 法は各種のスピルオーバーの過程を事例に即して同定した上で、主観的なアンケートをもって経済効果を金銭的な形で分析しようとする方法である。

一方、経済性効果の視点から外れるような社会的効果の分析についての講義も別枠で

行われる。日本では現在のところ経済性効果に焦点が置かれる傾向が強いが、社会的効果は経済性とは異なる側面から政策介入を正当化するものである。社会的効果としては環境、健康、安全、社会参加、犯罪抑制、文化継承、科学技術理解、防衛、雇用など様々な側面があり、特にその評価においては新しいアクター（社会グループ）を参加させる必要が指摘される。

（５）評価の利用

最後に評価の利用として、実際に意思決定に利用されやすい評価の行い方、評価結果の示し方とはどのようなものであるかが説明される。

そこでは、評価者（特に具体的調査を行う評価の専門家）と意思決定者の間で生じるコンフリクトはどのようなものであるかをまず説明し、評価の顧客は誰であるのか、評価レポートを受け取る人間が受け入れやすい提言とはどのようなものであるかを意識してレポートを作成する必要が簡単に説明される。

これら以外の内容としては、ケーススタディとして、英国 HEFCE が行っている大学の研究評価である Research Assessment Exercise、EU のフレームワークプログラムの評価のシステム、英国 DTI の評価システムが紹介される。後者 2 つについては、当該組織からゲスト講師を招いて講義が行われる。また、サセックス大学 SPRU からのスピナウトによるコンサルタント会社であるテクノポリス社の Erik Arnold 氏を招いて同社が行っている評価についての説明も行われる。

以上のように、PREST で行われている内容は、PREST の 1 年間のマスターコースの学生が受講生の大半を占めるために、研究開発の評価について基盤的な部分を幅広く紹介したものとなっている。各内容のより詳細な部分については、講義中に示される参考文献などを各人が独自に学習することになる。一方で PREST では、当然ながら、正規学生に対しては研究評価以外にも授業が行われているため、イノベーション研究や科学技術政策研究一般の内容、さらにはより一般のマネジメント理論や政策研究については別の講義で学習することになっている。そのため、それら背景的な知識についてはこの教育コースの中では教える必要がなく、教育コースは評価に直接関わる部分に比較的限定した内容となっている。しかしながら、本来は評価のシステムを形作る上で対象やそのマネジメントについての知識がある程度必要であることは、自明のことである。この点において、日本で研究評価の教育コースを開く場合には、受講生がこのような背景的な知識を有しているかを考慮して教育コースを形成すべきであると言える。

事例A-2. オランダ・トゥエンテ大学における研究評価の教育コース²

A-2.1 教育コースの概要

オランダのトゥエンテ大学においても、PRESTと同様に1週間ほどの教育コースが開講されている。このコースは12年前から開催されているものである。トゥエンテ大学の場合には、研究評価のみならず科学技術政策研究やSTS研究の第一人者であるArie Rip教授が中心となり教育コースを開講している。全体の講師は、トゥエンテ大学在籍の講師がRip教授を入れて2名のみであり、この他に、前述の英国マンチェスター大学PRESTのDr. Keenan、ドイツFraunhofer Institute Systems and Innovation ResearchのKuhlman教授、米国ジョージア工科大学のShapira教授といった外部講師を招いて構成されている(2002年の場合)。また、ケーススタディとしてEUやオランダVSNU(大学協会)からもゲスト講師を招いている。

トゥエンテ大学の教育コースは3日半という短い日数で行われる。だが、各日は朝9時から夜9時までという時間的に密なスケジュールになっているため、PRESTと比べてもほぼ同様の時間であると言える。2002年は3月4日～7日に開講された。受講料は1,495ユーロであった。

教育コースで教えられる内容は、大筋ではPRESTと変わらない。これは、そもそも欧州では1980年代から研究評価の研究者の間でシンポジウムや共同研究などによって密な情報交流が行われており、PRESTおよびトゥエンテ大学双方の教育内容はこれら欧州全体での議論(さらには米国の研究も含む)を背景として形成され、改善されているためであると言える。

A-2.2 教育コースの内容

トゥエンテ大学の教育コースの内容は表5.2に示すようなものである。コースは大きくは次のように区分されている。

- ・ 評価の社会的文脈
- ・ 評価のデザイン
- ・ 方法のカタログ
- ・ 政策のための評価

² 内容については、トゥエンテ大学でRip教授とともに教育コースを担当しているDr. van der Muelenにヒアリング調査を行った。

表2 トゥエンテ大学の教育コース内容 (2000年10月)

1日目	評価の社会的文脈	イントロダクション 評価の社会的文脈 コース概要
	評価のデザイン	評価のデザイン 演習1: 目的からデザインへ(ロジックチャートの演習と講評)
	方法のカタログ	事例: オランダの大学評価におけるピアレビュー
2日目	方法のカタログ (つづき)	方法のカタログ(概要) 中間的インパクトとしてのネットワーク 「社会的質」の評価 経済的インパクトの評価 事例: EUの研究評価 演習2: 評価の方法の総合的利用(演習と講評)
3日目	政策のための評価	評価データを用いた研究政策の改善 事例: 政策的介入の方法としての評価 因果関係の理解 演習3: 評価レポートの作成
4日目	政策のための評価 (つづき)	演習3(講評) コースのレビュー 将来の研究評価

(1) 評価の社会的文脈

最初の「評価の社会的文脈」は教育コースのイントロダクションであり、PREST 同様に評価の必要性を説明することで、なぜ評価を行うかという動機付けを再確認させるものである。この講義では、1980年以降の研究システムの変化により、研究評価がより複雑で困難な営みとなっていることが説明される。すなわち、1980年以前には競争的資金と委託研究との区別がある程度明確に線引きされていたために、それぞれピアレビューおよびプロジェクトの契約目標達成の分析という形で評価を行うことが可能であった。しかし1980年以降に戦略研究という新たなプログラム概念が生じ、より基礎的な研究開発行為についてもその社会経済的効果の中・長期的なスパンで分析することへの要求が生じてきたために、評価がより複雑なものとなってきている。

(2) 評価のデザイン

次に評価のデザインとしては、PRESTのDr. Keenanを講師として、上述のロジッ

クチャートの説明がなされる。実習の対象としては、2001年の場合にはやはり EU の EUREKA が用いられている。

(3) 方法のカタログ

評価の方法論としては社会的効果、経済的効果の分析手法が紹介される。科学技術的効果の分析手法の代表であるビブリオメトリクスについては数年前までは独立の時間割をくんでいたが、受講者の多くがその内容を既に知っている場合が多いため、独立の時間割ではなく他の説明の中に入れ込む形で簡単なものにしたという。一方、オランダでは社会的効果への注目が他国よりは強いという指摘がなされているが (van der Muelen and Rip 2000)、彼らの最近の研究成果を踏まえた分析手法の紹介がなされている。そこでは特に、新たなアクターを評価に取り込む必要が指摘され、そのようなシステム形成や、実際の環境改善や疾病治療への効果やその他の QOL への効果、規制の変化、国際的ポジショニングの変化などが評価の具体的内容として挙げられる。また、評価の示し方も社会的効果以外も含めた多様な項目をポートフォリオとして示す方法などが紹介される。

また、経済的効果についてはやはり PREST と同様に、スピナウトの概念や BETA 法が中心的に説明される。

トゥエンテ大学ではこの「方法のカタログ」の中で「中間的インパクトとしてのネットワーク」というタイトルの授業が開かれている点が PREST とは若干異なる点である。ここでは、プログラムの中間的なアウトプットとしてのネットワークへ焦点を置くことの必要性とネットワークをトレースする方法が説明される。これは、研究開発やイノベーションが、大学研究者や民間企業研究者などの共同研究や情報交流、さらには、それらに加えてユーザーとの交流をも通じて展開するものであるため、そのネットワークへ焦点を当てるのが研究実施のプロセスを考える上でも、顧客が誰でありいかに相互作用すべきかを考える上でも重要であることを認識するために必要な授業と考えられる。また、企業間や産学間、あるいは国家間での共同研究を誘因することを目的とするプログラムではネットワーク形成はより明示的に評価でも取り上げられるべきものである。この内容は、部分的には PREST でもスピナウト効果や行為のアディショナリティの説明、および、研究実施とその効果の関系の多様性の説明においてなされているものであるが、ネットワーク形成自体を特別な時間枠をとって教えるものではなかった。これは前述のように、PREST では別途イノベーション研究の授業が用意されている事も影響している可能性がある。

また、ピアレビューについてはオランダの VSNU (大学協会) の Steijn 氏を招いて、大学の研究評価におけるピアレビューの事例を紹介する。大学評価を事例として紹介して

いる点では PREST と同様であるが、オランダの大学評価は英国 RAE とは目的が異なり研究資金配分を行うものではないため、その方法は異なるものであることには注意が必要である。

2 日目の最後にはこれら方法論を踏まえて、英国 Centre for Urban and Regional Development Studies などを対象とした実習を行う。ここでは評価目的の設定と、それに適切な評価方法の選択が学習される。

(4) 政策のための評価

3 日目には Kuhlman 教授により、ISI が行った臨床研究センターの評価などを事例にしながら評価と政策との関係が説明される。評価は公的資金の使用法のチェックや目標達成の確認だけを目的とするのであれば政策的意義が薄いものとなる。政策やプログラムがどのように適切に展開され、今後もされうるかを評価で問うことが重要となる。ここでは評価対象となっている資金提供プログラムと実際に評価過程で同定された効果との因果関係の問題なども議論される。

また、これらを踏まえた上でいかに評価レポートを作成すべきかが、実習を含めて教えられる。ここでは、初日に扱った EUREKA が事例として再び用いられる。

最終日には、今後の評価システムで焦点となる問題についての Rip 教授の講義でコースは締めくくられる。

以上のように、基本的な部分は前節で紹介した PREST の教育コースと変わらない。欧州の研究者の間では研究評価のシステムや手法の現時点でのスタンダードについて、ある程度のコンセンサスが得られていることが伺える。ただし、社会的効果の分析やネットワーク形成の分析、サービスセクターの評価などは今後の研究評価の焦点となるものであるため、実際の評価や評価研究が進むにつれて内容は改善されていくものと思われる。

(大学評価・学位授与機構 林 隆之)

事例 A-3. 米国 ジョージア工科大学 (GIT) における教育コース

ジョージア工科大学 (Georgia Institute of Technology) の School of Public Policy³では、修士・博士課程における Areas of Concentration (集中分野) の1つとして Science and Technology (科学技術政策) が設定されており、この一環として研究評価に関連する講義を行っている。

修士・博士課程のカリキュラムは、コアカリキュラムと選択科目からなっており、選択科目で科学技術政策を選ぶことが出来る。科学技術政策関連の選択科目を表 3.1 に示す。これらの他にも、「Special Topics」として、研究評価関連の講義が行われる場合もある。

表 3.1 修士・博士課程の科学技術関連の選択科目 (PUBP は、科目コード番号)

選択科目 (科学技術政策)	
PUBP 640	科学、技術及び公共政策: 科学技術の支援、コントロール政策を含む科学技術と政府との関連の検討
PUBP 6402	研究政策とマネジメント: 研究政策およびマネジメントの課題の検討。公、民間および非営利組織の研究活動は、戦略的計画、資源配分、技術移転および研究評価の実践を検討する際に対比させられる。
PUBP 6414	技術的革新と政府の政策革新を促す連邦政府と州の政策。革新の源と刺激。大学と企業コンソーシアムの役割。革新政策の比較。技術政策の評価。
PUBP 6415	地方、技術および政策: 概念、問題、および地方開発、経済開発、産業の変化および技術政策と関係する政策の調査。
PUBP 6417	科学技術における重要観点: このコースは、学生の科学技術に関するクリティカルな思考、および市場と政治、社会との関係を刺激するよう努める。議論は、科学技術コミュニティの社会構成、科学技術における経済と政治力の役割などのトピックを含む。
PUBP 6421	大規模な社会技術システムの開発: このコースは、科学技術の社会、政治、文化的側面、及び国の規制政策や研究開発にどのような影響を与えるかを検討する。
PUBP 6608	技術のマネジメント: 外部環境
PUBP 6753	科学技術政策の比較
PUBP 6777	新興技術の分析: このコースは、技術のモニタリング、予見、評価手法の利用技術を開発する。さらに、新興技術領域の現状や見通しを検討する。

出典: GIT ホームページ⁴

³ <http://www.spp.gatech.edu/>

⁴ <http://www.spp.gatech.edu/academics/master.html>

訪問調査では、5名の先生方および2名の博士過程の学生から聞き取り調査を行った。以下にその概要を示す。

A-3.1 Bhaven N. Sampat 助教授

1) カリキュラム内容

Sampat 助教授が教えている、「科学技術政策研究の測定問題 (Measurement Issues in S&T Policy Research)」⁵というコースでは、科学技術が社会・経済的に与えるインパクトを、ありとあらゆるアプローチで検討している(表)。加えて、科学技術が社会に影響を及ぼすかを評価し、その評価を政策立案者がどのように役立てられるかについて研究している。特に、特許あるいは特許引用データの測定を行い、大学で行われている基礎研究が社会・経済的にどのような影響を与えているかに関する研究に力を入れている。

表 3.2 科学技術政策研究の測定問題

▶ 内容

このコースでは、科学技術の測定、科学技術の社会経済的影響の評価及び、科学技術政策変更の効果を評価するために、社会科学者に広く使用されているデータと判定・評価方法の概要について、講義を行う。

第一部では、社会科学での測定と因果説についての概要説明、科学技術政策における具体的な測定問題の検討、そして近年科学技術政策の立案者にとって、なぜ測定問題が重要になってきているのかについて議論する。

第二部では、ビブリオメトリックデータ、特許と特許引用データ、R&D データ、生産性データ、社会収益率データ、人的資源データなど、学者や科学技術政策の実務者に広く使われているデータを吟味する。具体的には、これらデータの入手方法、データから作成される指針、特定のデータや指針のメリットと限界について議論する。

第三部では、科学技術政策やその関連分野での近年の実例についての資料を幾つか選んで議論する。生徒の興味の範囲によっては変更の可能性もあるが、暫定的に我々が批評する調査は以下の通りとする：1) 科学の社会的・経済的影響の評価；2) 国際的技術普及の決定因評価；3) 地理が知識フローに与える影響の評価；4) Bayh-Dole 法の社会福祉に対する影響の評価；5) 国家改革システムの評価と比較。

第四部では、生徒は「セミナー」形式でコースレポートを提出する。このレポートに対して講師、クラスメート、(場合によっては) School of Public Policy の他の講師からコメント、批評、提案が提供される。

⁵ PUBP 8803 Special Topics での授業である。

▶ コース参加条件

このコースの条件は a) クラスへの出席 b) 各クラス前に指定された図書を読むこと c) クラス討論への参加 d) 判定レポート(referee report)2編 e) 最終レポートである。判定レポートは春休み前、最終レポートは最終クラスを締め切りとする。

このコースは学位論文のためのリサーチを補足するものであり、それに代わるものではない。そのため、それが科学技術政策研究の測定問題と何らかの係わり合いがあると私が納得するものであれば、判定レポートは自分で選んだ研究報告書についてでもよいし、最終レポートは自分で選んだトピックについてでもよい。最終レポートは研究プロポーザル、実証レポート(empirical paper)、批評文書レビューでもよい。

▶ 必須条件

なし。ただし統計のコースを事前に履修していると、講読が楽になる。

▶ コース概要

1. 序論

非実証データの原因の推測

科学技術政策の測定問題概要

2. 広く使用されているデータや指針についての概要

ビブリオメトリックデータ

パテントデータ

パテント引用データ

R&D データ

生産性データ

波及効果と社会収益の尺度

人的資源データ

3. 適用例

基礎科学の社会的・経済的影響の評価

国際的技術普及

地理と波及効果

Bayh-Dole 法の社会福祉に対する影響の評価

改革システムの比較

4. 生徒による発表

2) 受講者の卒業後の進路

最近の学生は、科学技術政策というよりも、環境政策に強い関心を持っており、その方面への就職希望者が多い。卒業生の中には、連邦政府や、ランドコーポレーションなどで科学技術政策を担当している人もいる。

3) 日本での研究評価教育プログラム策定に関する意見

研究評価のプログラム策定に関して、次の指摘を受けた。

「研究評価のプログラムを設ける場合には、3つの要素を考慮する必要があると考える。

一つ目は、科学技術の変化に関するプロセスを研究し、その変化が社会・経済的にどのような影響をもたらすのかを見ることである。例えば、歴史的な背景を見ることも重要であるし、ケーススタディーを行うことも重要であろう。異なる影響をもたらす研究が、どのような形で行われてきたかを、定量的に見てみるとよい。アプローチとしては、定性的・歴史的アプローチと、定量的・統計的なアプローチがあるが、両方からアプローチを行い評価するとよい。」

「二つ目は、統計学的な分析を行うことである。社会科学的数据の変化の要因を分析し、理解する。科学技術の評価の難しさは、知識や情報は目に見えるものではない、具体的なものではないという点である。そのため、相関関係的な結論を利用するために、統計学的な論理的手法を熟知することが重要である。指標 (Indicator) と概念 (Concept) を混乱することなく、分けて考えることが大切である。そして指標と概念では、概念が指標より重要である。」

「三つ目の要素としては政策立案者がどのような形で評価を行っているかの調査である。この調査には、特許データ分析、ビブリオメトリクス的な手法、研究開発のデータ分析、生産性のデータ分析などを利用する。

このような様々なデータを利用し、費用対効果がどうであるかなどを批判的な目で見たり、厳しいレビューを行うことも重要である。このような分野を学ぶ学生は、最近なされた評価の文献を常に読むことも重要であろう。」

「これら3つの要素を評価プログラムに取り入れることができれば、アメリカで行われているものよりも、優秀なプログラムができるだろう。アメリカではこれらの要素をすべて含むプログラムはない。

研究評価を行う際には、定量的情報と、定性的情報の両方を見て、そのプログラムが目的を達成できているかを見る必要があるが、米国でもまだそれはなされておらず、定量的なところに焦点を合わせる傾向がある。なぜならば、それが最も測定しやすいと思われているからである。しかしそれでは、ごく一部の小さい部分にしか焦点を合わせたことにならず、ゆがんだイメージしか得られないだろう。」

A-3.2 Barry Bozeman 教授及び Sooho Lee, Min-Wei Lin

1) 研究内容

Bozeman 教授は、DOE、NSF、NIH などの政府及び政府関係機関からの依頼によ

り、大学・政府の研究所およびそれらのプログラムの評価を行っている。

現在は NSF のグラントで、NSF のエンジニアリングリサーチセンターと科学センターの評価を行っている。また NIH のグラントで、NICHD (National Institute for Child Health and Development) の研究センターの評価を行っている。

その他、1000 人以上の科学者、エンジニアの履歴書を集めてコード化し、科学者、技術者のキャリアについて統計分析を行う「The Research Value Mapping Program」⁶を行っている。これは、世界で初めての試みである。

これらの研究には、リサーチアシスタントシップ (RA) 通じて、学生も参画している。RA とは、教授の研修を学生が補助する制度であり、学生は Bozeman 教授のもとで実際のプログラムに参加して、評価活動を経験できる。RA は授業より実践的であり、興味深い経験を得ることが出来る革新的なアプローチであり、調査テクニック (調査へのアプローチ、質問票の書き方、インターネットを用いた調査方法など) の修得や、実践を通して自分の研究テーマを発掘することが目的となっている。

2) リサーチアシスタントについて

プロジェクトを行う際に資金的余裕があれば、教授はリサーチアシスタントを雇うことができ、教授の手伝いをしてもらう。リサーチアシスタントには、学部、修士、博士課程のいずれの学生でもなることができ、RA を通じて実際のプログラムでの経験を得ることができる。

教授も学生に対して RA 制度を利用して進んで研究するように勧めている。学生は教授を助けるだけでなく、自分たちのアイデアを研究セミナーなどで発表することができる。自分の研究アイデアを追求していくことができる点が RA 制度のよい点であり、多くの場合大学院生は、プロジェクトに携わることで自分の研究テーマを見出し、卒業論文を書いているとのことである。

3) 受講者の卒業後の進路

博士課程の卒業者は、殆どが大学へ就職している。

A-3.3 Gordon Kingsley 助教授

1) カリキュラム内容

Kingsley 助教授は、プログラム評価に関して科学技術政策のクラスで教えている。全ての生徒が研究の設計から統計という基本的な技能を身につけるよう、一連の評価

⁶ <http://rvm.pp.gatech.edu/>

手法について教えている。

2) 受講対象者

修士課程、博士課程の生徒を対象としている。学部生に対しては、研究開発評価に関する授業はない。

GIT では、科学技術的な背景を持っている学生と人文・社会学の背景を持っている学生を、クラスに半数ずつ所属させ、意図的に混在したクラスを作っている。同じ背景を持つ学生を集めて授業を行うことは、授業内容を集中できるという面では効率的だが、混在させることにより、お互いから学ぶことができる、お互いの実態を知ることができるという点で効果的であると考えられている。

自然科学・工学系の学生と、社会科学・人文学系の学生を比較すると、例えば誤差に対する認識の仕方は大きく異なる。一般的に社会科学・人文学系のアプローチの方が、統計誤差は大きく出る傾向があり、それを自然科学・工学系の学生は理解する必要がある。自然科学・工学系の学生に実際の評価作業にあたってもらうことも将来的に評価の意義を理解する上で効果的である。

3) 受講者の卒業後の進路

修士課程を卒業した学生の40～45%が政府機関、40～45%がコンサルタント会社、10%程度がNPOに就職している。政府機関に就職した者は、科学技術政策を担当しているものと、環境政策を担当しているものに分かれる。これらの人々は、プログラム運営を行っている人の補佐を果たす、あるいはプログラムオフィサーとして実際の研究プロジェクトに対して支出を行っている部署で働く、もしくはプロジェクトの内容を評価する部署に属している。

博士号を取得した学生は、大学に残り教授の道を歩むことが多い。カンザス大学やコロンビア大学で教鞭を執っている卒業生がいる。また政府から委託され、研究評価を行っている卒業生もいる。

勤務先からの派遣等で博士課程を卒業した者は、元の所属していた機関に戻る人が多い。例えば、アトランタにあるCDC (Center for Disease Control) から来てPhDを取得した学生は、CDCに戻ったあと以前よりも高い地位に就いている。NSFから来ている学生もNSFに戻るようである。

4) 日本での研究評価教育プログラム策定に関する意見

研究評価の教育プログラムの策定に関して、次の指摘を受けた。

「受講の対象者のバックグラウンドにも注意すべきである。科学技術的な背景を持っているのか、公共政策的な背景を持っているのかで、研究評価に対するアプローチは

違うため、研修内容も変えた方がよいだろう。科学技術的な背景を持っている人に対しては、公共政策的な内容を、公共政策的な背景を持っている人には、科学技術的な内容を教えると良いだろう。

また、自然科学・工学系の教育を受けた人と、科学技術政策の教育を受けた人の考え方の違いを認識することも重要である。自然科学・工学系の人には、研究評価を懐疑的な目で見えており、研究評価の意義を理解するのに時間がかかる。しかしながら、将来的に研究プログラムを担当するポストに就いていることも多いため、これらの人に対して研究評価の意義、重要性を理解してもらう必要がある。また、実施される研究が社会全体に対してどのような影響をもたらすのかも理解してもらうことが必要である。

政府の研究開発プログラムの担当者は、より上級の職に就くに従って、研究が社会に与えるインパクトを連邦政府や議会に対して説明する必要性が生じてくる。そのため、その重要性を理解し、取り入れることで、報告がスムーズになることを理解してもらうことも重要である。さらに、研究開発の資金は国民の納税によって賄われていることも意識させるべきであろう。ややもすると、プログラム資金は天から降ってきたお金と思いがちである。」

A-3.4 Richard P. Barke 準教授

1) カリキュラム内容

Barke 準教授は、研究機関がどのような形で運営されているかという点に、特に関心を持っている。プロの研究者になるには、研究機関がどのような形で運営されているのかを理解する必要があるという考えである。従ってここ数年間は、研究政策、管理のコースを教えている。

表に、Barke 準教授が教えるコース概要について記す。

Barke 準教授のコースでは、以下の4つのアイデアを基本的に教えている。

一つ目は、研究機関が研究の対象となる優先課題が何であるのかを、どのように定めるかということである。

二つ目は、研究プログラムを実際にどのように実施するかである。これは、研究に対する支出がどのように、そしてどの程度成されているのか、人事はどう行われるのか、政府の政策がどのように影響するのか、研究が成功裡に終了したとき、その成果をどうするのかということである。

三つ目は、科学者が研究の成果・結果を、どのように伝達するかである。どのように発表するか、出版するか、あるいは専門誌に投稿するか業界紙にするか、どのように情報を分かち合うかなどである。そして技術者の場合は、ライセンス許諾の問題、

技術移転、特許はどのようにするのかという問題となる。

四つ目は、研究プロジェクトをどのように評価するかという、研究評価である。

そして、これらの4つの機能が、大学、産業界、政府の研究所でどのように実施されているかを見るのが、Barke 準教授のコースである。

表 3.3 研究政策とマネジメント⁷

▶ コースの目的

多くの政策、工学、科学コースでは、生徒にリサーチを行う方法を教えている。それに対してこのコースでは、研究機関がどのように運営されているか、そして生徒が自身のキャリアを通してアイデアや技術を適用するためにどのようにして自身の能力を最大限発揮することが出来るのかを中心に講義が行われる。このコースは全ての研究機関に共通する4つの主要なテーマ：プライオリティ設定(どの研究を実施するか?)、管理(適切な経済的・人的資源の獲得と使用)、活用(科学的コミュニケーションと技術移転/使用許可)、評価を主軸としている。2001 年春コースの主要テーマは「科学技術機関としての研究大学」とする予定だが、講読や事例、スピーカーなどを通してコースの範囲は個人セクターや政府研究機関にまで広げられる見込みである。

このコースはセミナーとして講義、ディスカッション、ゲストスピーカーにより行われる。戦略起案、資源配分、技術移転、研究評価といった研究方針や管理における具体的な課題についてより詳細に知ることが出来るよう、事例に基づいたケーススタディが行われる。生徒はグループで事例に関する問題点の解決策を考案し、創造性や現実性を試すため、その所見や勧告を実際の担当者と共に吟味する。講読とディスカッションにより、実際的な技術と批評的視点を兼ね備えることが出来る。

このコースの目的は研究機関が直面する課題を理解することである：

- －研究機関はどのようにして設立、運営、管理されているか？
- －研究プライオリティはどのようにして確立されているのか？
- －研究プロジェクトやプログラムを管理する際の主要な課題は何か？人材、予算組みと資金調達、法的・倫理的考慮など。
- －研究プロジェクトが新しい知識を生み出した後どうなるか？
- －研究者、研究プロジェクト、研究機関を評価することが可能な基準、指針、技術とは？

課題は研究機関の業務を知ることだけにとどまらない。このコースでは、なぜ研究所はそのような行いをするのか、また研究者や研究マネージャーが下した選択の倫理面についても検討し、生徒に現在実践されている行程の代案を考えてもらう。

⁷ PUBP 6402 の、選択科目である（表参照）。

▶ 宿題と成績

コースプロジェクトはケーススタディで、学期中に取り組み学期末には発表可能な状態で提出する。内容は地元の大学と提携している実際の研究機関についての調査で、以下について調べる：

1. 研究所の起源と使命
2. 教育、研究、その他の領域内でのその研究機関の位置付け
3. その研究機関の使命を同様の他研究機関のものと比較
4. 外部・内部管理
5. 内部管理構造
6. 内部組織構造
7. 職員：募集、選考、報酬：職務割当はどのように行われているか
8. 機関内の研究プログラム事例
9. 全体の研究ポートフォリオ
10. 課題設定と意思決定

コース中に何回か経過報告のためのクラスミーティングを行い、最終的には最後のクラスで詳細にわたる専門的なプレゼンテーションを行い、レポートを提出する。

▶ コース概要

- 1月8日 : コース紹介:「研究政策とマネジメント」と序論
- 1月10日・17日: 研究大学: 起源、歴史、役割、他の機関との関係、地域的・国際的背景、政治的・社会的背景
- 1月22日 : 研究大学の組織
- 1月24日 : 研究大学の管理
- 1月29日 : スピーカー
- 1月31日・2月5日: プライオリティ設定: 編成、基準、戦略プランと予測、参加者
- 2月7日 : 生徒によるレポート
- 2月12日・14日: 複数機能のバランス: 教育、研究、経済発展、サービス
- 2月19日 : スピーカー
- 2月21日・26日: 人事管理: 教授、スタッフ、生徒: 協力の形態
- 2月28日 : スピーカー (NAACP Julian Bond 氏、技術大学の社会的責任について)
- 3月12日・14日: 大学研究の管理: 資金調達
- 3月15日 : Henry Louis Gates (Harvard 大学)、Shirley Malcom (AAAS)、James Duderstadt (Michigan 大学)、Zell Miller (Georgia 州上院議員) による演説
- 3月19日 : プロジェクト管理: 研究の構造
- 3月21日 : 生徒によるレポート
- 3月26日・28日: 研究の活用 (コミュニケーション、使用許可、技術移転)
- 4月2日 : スピーカー
- 4月4日・9日: 研究の評価 (ピアレビュー、プロジェクトレビュー、プログラムレビュー)
- 4月11日 : 生徒によるレポート
- 4月16日・18日: 研究大学の将来
- 4月23日・25日: 生徒によるレポート
-

このコースでは、研究評価については3つの手法を用いて教えられている。一つは本や専門誌などの記事を読むこと。二つ目は、これらのトピックに関して経験を有しているゲストスピーカーの話聞くこと。例えば技術ライセンスの担当をしている大学や役所の担当者の話を聞き、質問を行うことである。書物などの論理的な内容と、実際に行われていることの両方を知ることが重要である。三つ目は、実際に研究を行っている研究所に対するケーススタディーを行うことである。この際重要なことは、研究評価を行うにあたっての評価基準が何であるのか、ということを理解することである。

従って、研究所がどのような目的を持って、どのような活動を行っているのかを理解し、そして予算上、人材的な制約はどのようなものがあるのかを理解する必要がある。よい評価を行うために不可欠なのは、組織そのものの評価を行うだけでなく、組織においてどのような環境が存在しているのかを理解し、それを評価することであるとの指摘である。

2) 受講者の卒業後の進路

昨年ケーススタディーを行った18人はすべてPhDコースの学生である。そのうち12人が公共政策を学んでおり、6人はエンジニアリングなどを専攻していた。公共政策を学ぶ12人中、博士号を取得したのはまだ2名であり、この2名はそのまま大学に残っている。

A-3.5 Susan E.Cozzens 教授

1) GITでのショートコースについて

GITでは現在、評価運営者を対象とした研究評価の短期コースを検討している最中である。内容的にはTwente大学と同様の内容を想定しており、例えばある年はTwente大学で教え、ある年はGITで教えるという形を取りたいということである。

コースの対象者は、中堅クラスの研究管理・運営業務を行っている人たちである。期間は1週間程度を想定しており、2003年10月に開講することを目指している。コースの名称はInstituteやWorkshopとしたいとのことである。

現在政府機関で研究評価を行っている人の殆どは、研究評価の教育訓練を受けていない人であり、自然科学・工学系の専門分野の人が多く。そのため、彼らを対象としたコースを開講したいとのことである。

以下では、Cozzens教授が考える、短期コースの設計概念について示す。

▶ Cozzens 教授の短期コース設計概念

短期コースを設計する場合、研究評価の文脈がどのようなものかというところから出発する。なぜ研究プログラムの評価が必要なのか、外部からどのような圧力がかかっているか、そして内部的に研究評価を行う必要性はなんなのか、という点を示し、なぜ研究評価が必要なのか、なぜ評価をしたいと思っているのかを考えてもらうことが必要である。

また、いくつかの事例を用いた教育を行う。事例には簡単な事例と複雑な事例を含める。

簡単な事例としては、1つのプログラムを外部パネルに評価してもらうという事例である。ここで重要なのは、誰が、あるいはどこが評価を求めているか、評価結果をどのように利用するか、プログラムの目標は明確であるか、評価がその目標を明確にしうるか、という点を明らかにすることである。

これらを明らかにする際には、例えば次の点が問題として浮かび上がる。誰が評価するのか、評価する人を選ぶのが誰か、パネルのメンバー構成としてプログラムを理解している人とプログラムに距離を置いている人のバランスをどうするか、パネルにどれだけの資料を前もって提供するか、資料の形態は要約文か生データか、資料を実際の評価にどのように利用するのか、などを検討しなければならない。その他にも、レビューにどの程度の日数を費やすか、パネルはプログラムの中身を知っている人からヒアリングをするのか、それとも資料に目を通すだけなのか、報告書を誰がどのように、どの程度の量でまとめるか、報告書を受け取った者はそれをどのように活用するか、そして報告書は最終的に誰に渡されるのか、問題点を発見した場合や、助言を行った場合、追跡を行っていくのかどうかなど、考慮すべき点は沢山ある。

パネルでの評価が研究評価活動の典型的な形式なので、短期コースの場合、パネル評価でもこれだけ検討する点があるということ、参加者に認識してもらうことが大切である。

以上のような点を、参加者に質問して回答をもらい、お互いの答えを比較する、議論するというのもよいだろう。

複雑な評価事例を用いることも、データの収集や利用、外部委託の際に考慮すること、プログラムの結果や成果の測定方法などを考えてもらうために必要である。複雑な事例の場合には、基礎研究と戦略的な研究では評価手法が異なるため、いくつかの事例を取り上げる必要がある。

受講者が、研究評価の過程においてどのような測定方法が用いられているかを一覧表にし、それぞれの短所と長所を理解出来るようになればよい。

最後に、戦略的な施策の立案を行うときに、評価結果をどのように役立てるか

などについて議論を行うのがよい。

短期ではなく長期のコースを持つことができれば、自国でどのような研究開発システムが存在しているのか、革新的なプロセスとは何か、科学的な創造性の意義、科学技術をいかに奨励し守っていくか、ということを読んで欲しい。また、プログラムがどのような成果を上げているのかを定期的にモニタリングし、進捗状況を把握するためのパフォーマンスインディケータについても示せばよい。

2) 受講者の卒業後の進路

GIT の修士の卒業生は、多くが政府の省庁に入省し、政策立案や企画部門にいる。また何人かは民間に就職している。博士号取得者は、まだ2人しか卒業していないが、1人は環境政策を専門として大学に就職、もう一人は CDC からの派遣であり、CDC に戻ったとのことである。

3) 日本での研究評価教育プログラム策定に関する意見

研究評価の教育プログラム策定に関して、以下の指摘を受けた。

「米国で評価を担当する評価運営者は、評価の重要性は理解しているものの、どこから手を付けていいのかわからない人も多い。また、評価の内容についても、簡単なようであり、複雑なようであるという印象を持っている。そのため研修の際には、評価の概要を明らかにし、どのようにアプローチしたらよいかを提示することが重要だろう。」

(本調査と提携して行われた、(株)三菱総合研究所 岩崎裕典氏の調査による)

事例 A-4. AAAS 科学技術政策フェローのオリエンテーション・プログラム

AAAS (American Association for the Advancement of Science : 全米科学振興協会) 科学技術政策フェローシップは、全米各地の大学で科学技術分野の博士号または同等の学位を最近取得した若手の科学者が、ワシントン DC の科学技術政策の立案や意思決定を補助する目的で本格的には 1973 年に開始された。これまで 1200 人の科学者や技術者がこのフェローシップを経験している。

志願者はアメリカ市民でなければならない。それぞれの科学技術情報や専門的知見を連邦議員や連邦政府に提供しつつ、彼ら/彼女ら自身も、政府がどのような仕組みで動くのかを学びつつ、科学技術の社会への影響を学ぶことができる。

フェローは、連邦議会、全米科学財団 (NSF)、国立衛生研究所 (NIH)、国務省、国防総省、米国国際開発庁 (USAID)、環境保護庁 (EPA)、農務省、食品・医薬局 (FDA)、司法省、エネルギー省などのオフィスに配属される。

このフェローシップには現在では毎年、約 70 人のフェローが選抜され、1 年間に 5 万 5000 ドルの給費が支払われる。フェローシップは 9 月 1 日から始まり、それぞれの省庁に 1 年間派遣される前に、フェロー全員は 2 週間のオリエンテーションに参加する。また、年間を通じて科学技術と公共政策に関するセミナー、勉強会、パーティなど様々なイベントが頻繁に開催され、フェローの交流・情報交換の場となっている。

AAAS フェローシップの種類と業務内容

フェローシップには、2002 年から 2003 年にかけては下記の 10 種類がある。

▶ Roger Revelle Fellows Program in Global Stewardship

フェローは 1 年間、国内ないし国際的な環境政策の分野で、議会や関連する行政府機関もしくは他の政策コミュニティ内の組織で働く。応募者は生物科学、物理科学、社会科学のいずれかの分野の Ph.D を持ち、かつ最低 3 年間の学位取得後の専門的経験のあるアメリカ市民でなければならない。年俸は 5 万 5 千ドル。

▶ Congressional Fellows Program

フェローは、科学的・技術的インプットが必要な立法・政策分野における特別補佐としてキャピトル・ヒルで議会や議会委員会のメンバーとともに 1 年間過ごす。募集人員は二人、年俸は 5 万 5 千ドル。

▶ AAAS/NTI Fellows Program in Global Security

フェローは、公共衛生および医学に関する専門知識を応用して、1 年間生物兵器やバイオテロ拡散防止、連邦対応計画に関係する問題について働く。配属先は議会および行政府内の幅広い可能な選択肢の中から選べる。応募者は米国民であり、生物科学もしくは公共衛生あるいは公共衛生関連分野の医学博士号、もしくは DVM, Ph.D を有していなければならない。年俸は 5 万 5 千ドル。

▶ **Diplomacy Fellows Program**

フェローは、国務省の外交政策部門、もしくは国際開発庁（U.S. Agency for International Development）の国際開発部門、農務省の国際経済・農業開発部門、国立保健研究所フォガーティ国際センター（Fogarty International Center of National Institute of Health）の国際保健部門のいずれかで、科学・技術分野における国際問題に関して1年間働く。国務省は最大6人、国際開発庁は約12人、農務省のプログラムは一人、国立保健研究所は一人を採用予定。年俸は5万5千ドル。

▶ **Risk Policy Fellows Program**

フェローは、農務省もしくは食品医薬品局（U.S. Food and Drug Administration）のいずれかで1年間働き、リスクアセスメント/リスクマネジメントの衛生的・経済的・環境的側面に関連する問題について、科学的・技術的インプットを提供する。3~5人を採用予定。年俸は5万5千ドル。

▶ **Defense Policy Fellows Program**

フェローは、1年間国防政策や技術応用、国防システム分析、プログラム監視・管理に関連する問題に関して働く。配属先は、国防次官調達・技術・ロジスティクス各局（offices of the Under Secretary of Defense for Acquisition, Technology and Logistics）もしくは、陸軍研究課（Army Research Office）、防衛脅威低減局（Defense Threat Reduction Agency）。配属には、省庁横断的活動や議会活動、国際的活動が含まれる可能性がある。3人前後を採用予定。年俸は5万5千ドル。

▶ **Environmental Fellows Program**

フェローは、ワシントンD.C.の環境保護庁（U.S. Environmental Protection Agency）本部で科学・政策・環境に関連するプロジェクトに従事する。このプロジェクトには、リスクアセスメントに関するものも含む。10人を採用予定。年俸は5万5千ドル。

▶ **AAAS/NSF Science and Engineering Fellows Program**

フェローは1年間、全米科学基金（NSF: National Science Foundation）で過ごし、NSFがどのように科学助成を行うか学ぶとともに、基礎科学と工学研究・教育を支援するNSFのミッションに関連する問題について科学的・工学的・教育的インプットを提供する。年俸は5万5千ドル。

▶ **AAAS/NIH Science Policy Fellows Program**

フェローは1年間、国立保健研究所（NIH: National Institute of Health）で医学研究の現場に影響を与える政策の分析・開発・導入について学ぶ。また、フェローは、研究活動や研究トレーニング活動、コミュニケーション活動を支援・管理・調整するためにNIHが用いるプロセスとメカニズムについて理解を深める機会を提供される。年俸は5万5千ドル。

▶ **AAAS Science, Justice, and Public Policy Fellows Program**

フェローは司法省（DOJ: U.S. Department of Justice）で1年間過ごし、以下の司法省内組織のいずれかに対して技術上・研究上の支援を提供する。環境・自然資源部（the Environment and Natural Resources Division）および、連邦捜査局（FBI）、司法管理部（Justice Management

Division), 政策開発課 (the Office of Policy Development), その他上級管理局。フェローは、フェロー自身の関心や特定部局の必要に応じて司法省のほかの部局に所属される可能性もある。年俸は5万5千ドル

(WEB 情報を (財) 政策科学研究所がとりまとめたもの)

本節では、以下に私 (本調査協力者 下斗米一郎氏) が特別参加した 2001 年 9 月 5 日から 20 日まで全 11 日間の日程で行われた、AAAS 科学技術政策フェローのオリエンテーション・プログラムの概略について述べたい。

A-4.1 オリエンテーション・プログラムの概略

◇ 第1日目：9月5日 (水)

初日は、ワシントン DC のダウンタウンのほぼ中心部に位置する AAAS の建物内の講堂にて、AAAS 科学技術政策フェローシップ・プログラムのディレクターのクラウダ・スタージスさんらから、オリエンテーションと 1 年間のフェローシップの概略について説明を受けた。また、経験豊富な政府コンサルタントのマーク・タリスマン氏から、アメリカ政府の歴史的特徴について、概略説明を受けた。また、プログラム・スタッフや今年度のフェローの自己紹介が和やかな雰囲気で行われた。

夕方には、前年度までの AAAS 科学技術政策フェローで引き続きワシントンでスタッフとして働いている OB/OG と懇談する機会が設けられた。その後、歓迎パーティが行われ、新年度も更新した AAAS フェロー、AAAS 科学技術政策フェローのスポンサーになっている各学界の代表、フェローの受入先となるワシントンの政府機関やオフィスの代表やスタッフらが参加し、大いに活況を呈した。

◇ 第2日目：9月6日 (木)

2 日目の午前中は、「公共政策のどこに科学が適するか? ——国内と国際的な視点から——」というテーマで、ワシントンの科学技術政策の実務家・有識者 4 人が AAAS フェローの前で討論を行った。討論者は、科学技術と国際問題に関してワシントン屈指のシンクタンクである戦略国際問題研究所 (CSIS)・科学技術政策プログラムの上級アソシエイト (現在、私の上司でもある) の Anne Solomon さん、AAAS 国際プログラム・ディレクターの Richard W. Getzinger さん、AAAS 科学技術政策プログラム・ディレクターの Albert H. Teich さん、ジョージ・メイソン大学の Christopher Hill 教授である。

また、「科学者と政策立案者の世界観の相違——双方のコミュニケーションの結果」と題し、AAAS 科学技術政策プログラムのアソシエイト・ディレクターの Stephen D. Nelson さんが日頃の実務の様子を説明した。

午後には、国家の経済政策についてのレクチャーが行われた。経済政策についてワシントン屈指のシンクタンクであるブルッキングス研究所の上級フェローの Barry Bosworth 氏と国際経済研究所 (IIE) の Edwin M. Truman さんが講師となり、多くの公共政策の必須知識である経済政策のしくみと働きについて講義した。

その後、「連邦議会で要求される倫理」、「行政府で要求される倫理」と 2 つのグループに分かれ、それぞれの職業環境で何が許され、何が許されないのか、避けなければならない

状況とはどういうものか、といったフェローにとって必要不可欠な職業倫理教育が行われた。

「連邦議会で要求される倫理」については、下院「公的行為基準 (Standards of Official Conduct)」委員会の John Vargo さん、上院 Select Committee on Ethics の Beth Ryan さんが担当した。「行政府で要求される倫理」については、連邦政府倫理局 (Office of Government Ethics) の James O'Sullivan さんが担当した。

◇ 第3日目：9月7日（金）

3日目の朝には、「連邦予算手続きの紹介：予算を理解しなくては、政策過程を理解できない」というテーマで、AAAS 研究開発予算プロジェクト・ディレクターの Kei Koizumi 氏から講義が行われた。Koizumi さんは日系2世で、長年、研究開発予算分析を行ってきた Albert Teich さんから引き継いでおり、この講義はオリエンテーション・プログラムの中でも参加者から好評を得ていた。

引き続き、「予算決議過程：予算の意思決定に関する政治」というタイトルで、SLR 予算と立法コンサルタントの Shirley L. Ruhe さんがコーディネートしながら、参加者が実際に予算決議過程をシュミレーションする双方向ワークショップが行われた。

午後からは、「科学技術政策 53 年の回顧」と題し、立法府と行政府の高官ポストを歴任してきたカリフォルニア大学サンタ・クルーズ校の William G. Wells, Jr. 教授が講義を行った。

◇ 第4日目：9月10日（月）

この日の朝は、アイゼンハワー行政オフィス・ビルに集合し、ホワイトハウス科学技術政策の首席スタッフの Richard Russell さんとの懇談会が行われた。その後、「大統領への科学アドバイスの歴史」と題し、コロンビア大学 Heymann Center for the Humanities の Bruce L.R. Smith 教授からレクチャーが行われた。

ランチは、「大統領府について」と題し、メリーランド大学 James MacGregor Burns Academy of Leadership の Georgia Sorenson 教授との自由闊達な懇談形式のランチとなった。

午後は、「連邦議会の観点」と題し、連邦議会リサーチ・サービスの Walter Oleszek さんが講義を行った。その後、「立法過程」について、Sherwood Boehlert 下院議員（共和党・ニューヨーク州選出）オフィスの立法ディレクターの David Goldston さんがレクチャーを行った。

◇ 第5日目：9月11日（火）

この日は静かな朝で始まった。いつも通り、8時30分頃、AAAS 建物内にフェローが集まり、プログラム前の気軽な立食朝食会が行われた。そして、この日は9時から、「立法府トラック」、「国際トラック」、「行政府トラック」の3グループに分かれた。

私は、AAAS 講堂で行われた行政府トラックに参加した。「科学者のいる省庁におけるサイエンス・フェローの役割」と題し、環境保護庁 (EPA) の Michael Slimak さんと全米科学財団 (NSF) の Joanne Tornow さんの2人が講義をしている最中、突然、AAAS スタッフが壇上に上がり、「突然ですが、緊急事態が発生しました。」と緊張した表情でプログラムを中断した。「現在、国家非常事態となった模様です。」と耳を疑うような言葉を発し、講堂の大きなスクリーンに CNN の生映像がリアルタイムで映し出された。ニューヨークの世界貿易センターの1棟から煙が燃え上がっている。「火事か？」と思っていたところ、も

う1棟のビルに飛行機が激突した。「何てことだ！」会場からはざわめきや悲鳴が上がった。

ニューヨークだけの事態かと思いきや、近隣のペンタゴンにも激突された映像が映し出された。「アメリカが攻撃を受けている」ということをフェローたちは次第に理解し始めた。中には、家族や友人が犠牲になったかもしれないと泣き出すフェローもいた。そして、AAASスタッフが「皆さん、きょうのプログラムは中断します。一刻も早くダウタウンから避難して下さい。明日以降の予定は明朝、こちらに電話し、どうなるか確認して下さい。」と何度も声をかけ、みな一斉に建物を後にした。

ワシントンのダウタウンには、人や車があふれ渋滞状態だ。皆が郊外へ向けて一斉に足を運んでいた。パトカーや消防車のサイレンがあちこちで鳴り響き、混乱状態となっていた。

◆ 第6日目：9月12日（水）

早朝、AAAS に電話で確認したところ、午前中に国務省で行われる予定だったプログラムは状況が状況なだけに中止となったとのこと。この日は、実務家の方々のプログラムは中断され、「連邦政府の官僚組織の性質」と題したジョージタウン大学の Mark Rom 教授の講義のみ行われた。

◆ 第7日目：9月13日（木）

この日の午前中は、「立法府トラック」、「国際トラック」、「行政府トラック」の3グループに分かれた。

「立法府トラック」の参加者は、レイバーン下院ビルに集合し、まず「第107回連邦議会の展望」と「連邦議員事務所はどのように機能するのか」と題し、上下両院の連邦議員スタッフ9人（うち5人はAAASフェローOB/OG）が、共和党と民主党からバランスよく講義を行った。

「国際トラック」の参加者は、AAASのビルに集合し、まず「NGO（非政府組織）の役割」と題し、Africa Wildlife FoundationのKatie Frohardtさんと、AfricareのYolonda Richardsonさんが講義した。その後、「グローバル問題における米国のリーダーシップの役割：有用か、悪用か」と題し、ジョージタウン大学のCharles Weiss教授が、外交と科学技術がクロスした問題を提起した。

「行政府トラック」の参加者は、AAASのビルに集合し、「連邦アドバイザリー委員会」について、U.S. General Services AdministrationのJames Deanさん、全米科学財団（NSF）のMarta Cehelskyさん、米国医薬品局（FDA）のLinda Shermanさんが講義した。米国には約1000の連邦アドバイザリー委員会が存在し、科学と政策問題を含め、60の連邦省庁にアドバイスを提供している。「誰が、28000人の委員会メンバーとこれらの委員会に使用される18億ドルの資金を監視するのか？これらの委員会がAAASフェローのこの1年にどのようなインパクトをもたらすのか？」といった説明が行われた。その後、「政府のパフォーマンスと結果法（GPRA）」について、行政管理予算局（Office of Management and Budget）のDavid Radzanowskiさん、全米科学財団（NSF）のLoretta Hoplinsさんが説明した。

お昼には、すべてのフェローがフェニックス・パーク・ホテルに集合し、Rush Holt下院議員（民主党ニュージャージー州選出、AAAS科学技術政策フェローOB）とのランチが行われた。Holt議員自身も科学者で、AAAS科学技術政策フェローの経験、科学者であったバックグラウンドがそのように議員活動と関係しているか、などについてレクチャー

していた。

午後は、「米国外交政策：インタラクティブ・ワークショップ」と題し、リオにおける「環境と持続可能な開発」国際会議に米国政府代表として参加するシミュレーションを行った。このワークショップを通じ、フェロー自身が実際の政策担当者として、大きな国際会議に向けて準備することで、科学と外交政策、政治・経済の共通領域を理解することに主眼がおかれた。

フェローは各 10 人程度のグループに分かれ、どういったポジションの人間を代表として派遣すべきか、この会議で米国が達成すべきこと、どの程度の重要性をこの会議に米国は置くべきか、といった個別の論点がグループごとに提示された。私が参加したグループでは、「リオ国際会議の準備や会議そのものに連邦議会を巻き込むべきか？もし、そうならば、誰を、なぜ、どのように、いつ？」という論点が与えられ、グループ内で討議の上、結論が出された。グループごとの討議の後、全体総括に入り、各グループの代表がそれぞれの論点の結論／推薦政策を発表し質疑応答が行われた。

講義が多いオリエンテーション・プログラムの中で、このインタラクティブ・ワークショップは、参加者がそれまでに学んだ知識を総動員し、主体的に考え、議論し、結論を導くことから参加者から好評を得、もっとワークショップをオリエンテーション・プログラムの中に取り入れて欲しいという意見も聞かれた。

◇ 第 8 日目：9 月 14 日（金）

午前中は、連邦議会議事堂で「ワシントンにおけるロビー活動」と題し、実際にロビイストたちからロビー活動について説明を受けた。その後、ランチでは「エネルギー・気候の挑戦、そしてわれわれは何をすべきか」と題したレクチャーが、ハーバード大学ケネディ・スクールの John Holdren 教授から行われるはずだったが、テロ発生の影響で中止された。

午後はナショナル・プレス・クラブにて、「司法における科学」と題し、連邦裁判センターの科学的証拠プロジェクト・ディレクターの Joe Cecil さんから講義が行われた。裁判において科学的証拠をいかに集め、科学者がどういう役割を担うのか説明を受けた。

◇ 第 9 日目：9 月 17 日（月）

この日の午前中は、「国防総省における科学技術」と題し、ペンタゴンで講義を受ける予定だったが、テロの影響で中止された。

午後は、全米科学アカデミー（NAS）で、NAS プレジデントの Bruce Alberts 氏と、NAS 「議会と政府活動局」ディレクターの Jim Jenson 氏から、NAS の活動概略の説明を受けた。その後、「情報と誠実さをもって官僚組織の中で、どう影響を与えるか」という題で、国務省の Jonathan Margolis さん（1991-92 AAAS 国務省外交フェロー）が講義した。

◇ 第 10 日目：9 月 19 日（水）

この日の午前中は、まず各学問分野の視点から最近の主要課題についてブリーフィングが行われた。

「社会科学」分野では、下院「教育と労働力」委員会の Ruth J. Friedman さん（2000-01 AAAS 連邦議会フェロー）、アメリカ心理学界の Heather O'Beirne Kelly さん、米国国際開発庁の Mariott Pratt（1995-97 AAAS 外交フェロー）さんが担当した。

「バイオ医学分野」では、国立衛生研究所（NIH）の Kathleen Michaels さん（1992-94

AAAS 外交フェロー)、National Research Council の Laura Sheahan さん(2000-01 AAAS 連邦議会フェロー)、Henry A. Waxman 下院事務所 (民主党・カリフォルニア州選出) の Tim Westmoreland さんが担当した。

「環境・農業分野」では、農務省の Deborah L. Sheely さん (1995-97 AAAS 外交フェロー)、国際開発庁(USAID)の Barbara A. Best さん (1997-99 AAAS 外交フェロー)、環境保護庁 (EPA) の Stuart Lehman さん、James M. Jeffords 上院議員事務所 (無所属・ヴァーモント州) の Kathryn Parker さん (2000-01 AAAS 連邦議会フェロー)

「物理学分野」では、Office of Energy Efficiency and Renewable Energy の Allan R. Hoffman さん (1974-75 AAAS 連邦議会フェロー)、空軍科学研究所の Joan Fuller さん (1997-99 AAAS 防衛政策フェロー)、上院「武器サービス」委員会の Arun Seraphin さん (1997-99 AAAS 連邦議会フェロー) さんが担当した。

「工学技術分野」では、Institute of Electrical and Electronics Engineers の Chris Brantley さん、Tony P. Hall 下院議員事務所 (民主党・オハイオ州選出) の Julie Pollitt さん (2000-01 AAAS 連邦議会フェロー)、米国国際開発庁の John Ingersoll さん (2000-02 AAAS 外交フェロー) が担当した。

その後、「ワシントンのシンクタンク」と題し、ブルッキングス研究所の Kent Weaver さんが講義を行った。シンクタンクの役割や特色、運営方法などを説明した。

午後からは、「政治と科学が異なったアジェンダをもつ時～食品・医薬局 (FDA) とタバコ規制の事例～」と題し、アメリカ心理学界の Daniel Dodgen さん (1996-67 AAAS 連邦議会フェロー) が司会となり、規制する側と規制される側のそれぞれ異なる視点をもったスピーカーが討論した。

その後、「ワシントンで科学と科学政策ニュースを報道すること」と題し、ニューヨーク・タイムスの Cornelia Dean さん、下院「運輸」委員会の報道担当の Jim Berard さん、National Public Radio の Richard Harris さん、国家航空宇宙局 (National Aeronautics and Space Administration) の David Steitz さんがスピーカーとなり、日頃の報道の様子や気をつけるべき点などについて話した。全米科学財団 (NSF) の Curt Suplee さんが司会を務めた。

夕方から、連邦議会フェローに関しては、連邦議会議事堂で、AAAS フェローを今年度スタッフとして採用することに関心のある連邦議員や議員事務所スタッフなどとのレセプション・パーティが行われた。

この他のフェローは、ユニオン駅近くのビール・バーでパーティを行った。

◇ 第 11 日目：9 月 20 日 (木)

この日は、連邦議会フェロー、米国国際開発庁フェロー、環境フェローなど、グループごとのミーティングとなり、各フェローの今年度の配属先が決定された。

(松下政経塾 下斗米一郎)

事例 B 評価人材の確保・活用に関する事例

B-1 米国 NIH におけるレビュー体制と評価人材に関する議論

B-2 欧州の行政機関等の評価とその人材育成の取り組み

B-3 EU における科学技術ピアの公募・選定・行動規則等

事例 B-1. 米国 NIH におけるレビュー体制と評価人材に関する議論

米国 NIH（国立公衆衛生院：National Institutes of Health）は、研究助成審査における二段階のレビュー・システムで知られているが、レビュー体制についての継続的な見直しが行われ、近年も調査や議論が進められている。そこでは、評価人材に関わる問題も扱われており、例えば、運営責任者（SRA：Scientific Review Administrator）の役割やピアレビューの資質や構成、適正負担の問題、ある種のレビューに科学技術ピア以外の評価者を加える必要性とその人選・活用のための配慮の問題などが論じられているので事例として取り上げる。

B-1.1 グラント審査の特徴—デュアル・レビュー・システム

NIH のグラント審査（対象はほとんどが基礎研究であり専ら事前審査）は、一般に 2 段階審査（デュアルレビュー・システム）でなされていることが知られている。第 1 段階のレビューでは、研究内容のみを審議する Scientific Merit Review と呼ばれるピアレビューがなされている。これは専門家を活用するもので、学会の審査活動とほとんど類似のものである。第 2 段階は、第 1 段階の科学的メリットの評価を踏まえ、NIH で実際に助成にあたる担当研究所のプログラム関連性などを反映させ、課題に実際に支出することの適否の政策的配慮 Second Level of Review を行う審査である。Advisory Council & Board により評価される。

プログラム/プロジェクトやセンターなどは Institute Initial Review Group が審査する体制となっている。

NIH ピアレビュー活動を統御する規則や規制についての公式文書は非常に多くあり、透明性は高い。CSR（Center for Scientific Review：科学レビューセンター、旧研究グラント局）に対する申請者への解説文書では、プロセスに沿った評価システムと評価業務の概要がよく記載されているので、B-1-資料 1 にまとめた。公正で応募者支援型のピアレビューを実現するような評価人材マネジメント（利害衝突の回避、秘密保持、科学上の不正があった場合の措置）に留意していることが分かる。

B-1.2 評価の運営人材の役割とレビューアの選考

第一段階のピアレビューは CSR が運営するが、その Study Section メンバー（レビューア）の人選は、NIH の職員である SRA とその Study Section の委員長が行う。

The Public Health Service (PHS) Act は、NIH 所長または定められた連邦職員（federal

designees) に対して、科学レビューグループを設置し、そのメンバーを任命する権限を与えている。定められた連邦職員とは SRA であって、SRA は Study Section のメンバーのリクルートと、彼らに対する申請書の割当について責任をもつ。SRA は、現実には殆ど博士で研究経験があり、学会でも有数の専門知識と権威を備えている。連邦政府レベルでは SRA の設置は、Federal Advisory Committee ACT に根拠があり、Designated Federal Official (DFO) と称される。SRA は、SRG (Scientific Review Group) の勧告に意図的に影響を与えることは禁じられるが、SRG の会合の設定に関わる作業に責任をもつと同時に、SRG が科学的に妥当な勧告に到達できるよう努力することが義務付けられている。

第 2 段階の Advisory Council/Board では、プログラムスタッフ/プログラムディレクター (PS/PD) が評価を運営する。

(注) “Reviewer Assignment to Applications,” Peer Review Notes, Sep. 2000 参照
study section の委員長には、SRA と共同でレビューアを選定する役割がある。メンバー構成の最終的責任は SRA にあるが、そのプロセスに積極的に委員長が関与するように Guidelines for Study Section Chairs では定めている。

参考 “Guidelines for Study Section Chairs,” CSR, 18th July, 2001
(<http://www.drg.nih.gov/events/guidelineschairs.htm>)

SRA によるレビューアの選考には基準がある。Guidelines for Study Section Chairs および Role of the SRA-A Quick Overview では、次のように述べている。

- a. 科学的卓越性 (グラントや公刊記録によって明らかに示されるもの)。
- b. 科学共同体における尊敬。
- c. 経験の幅広さ。
- d. レビューにおける公正さと中立性。
- e. 会合におけるプレゼンテーションの明快さと参加の質。

また、同時にメンバー構成に関しては、SRA が次の点について考慮していることに留意してほしいとの注意書きがある。

- a. メンバーが、必要とされる専門知識・技能の全体をカバーしているか。
- b. ジェンダーや民族、地理的位置に関する多様性。

(注) “Guidelines for Study Section Chairs,” CSR, 18th July, 2001.
(<http://www.drg.nih.gov/events/guidelineschairs.htm>)

“Role of the SRA-A Quick Overview,” CSR, 19th July, 2002.
(<http://www.drg.nih.gov/events/rolesra.htm>)

ピアレビュー・グループのメンバーの選定に際しては、遵守すべき倫理綱領が定められている。Peer Review Notes には、次の倫理綱領の引用がある。

The Code of Federal Regulations 42 CFR ch. 1, 52h.5 Conflict of Interest(10-1-97 Edition)

(注) “Update on Review Policy and Procedures,” Peer Review Notes, May 2001.

B-1.3 評価マネジメント人材に対する教育プログラム

NIH の評価システムのマネジメントに関与する専門スタッフ (professional staff) である SRA と the Program Official (or Project Officer) は、Extramural Science Administrators (ESA) と称されるが、全員に、NIH 全体の知識や情報を共有するようにトレーニング・プログラムが整えられている (NIH Record-2/06/2001—Training Plan Redeveloped for Extramural

Administrators 参照)。

第一に、2年未満の経験をもつ ESA にはコースワークがある。新人には NIH の所外に開かれた活動の概要を知る「オリエンテーション」が義務づけられており、OEP (the Office of Extramural Programs) が提供している。過去2年以内に初めて職務を経験した ESA (助成や契約のマネジメントや支援スタッフも) に対しては、知識を拡げスキルの更新をする毎週半日1ヶ月の「コア・カリキュラム・トレーニング」があり、ESA としての職務パフォーマンスの基盤を掘りさげることが出来るように、プログラムやレビューのマネジメントにおいて直面する複雑な 이슈 が扱われる。(http://odoerdb2.od.nih.gov/oer/training/esa/esa_cores.htm 参照)

第二に、それ以上の経験を持つ人材に対する継続教育活動と、さらに限定されたメンバーに対する高度教育プログラムがある。「継続教育」は、NIH の所外に開かれた助成等の活動に関する政策や手続きに関する最新知識を確保させるものである。ESA は、OEP が提供する特別イシュー・トレーニング、STEP (Staff Training in Extramural Programs) フォーラムないしモジュールへの出席、その他のしかるべき機関のトレーニングのうち、少なくとも2つを受講することを求められている。また、高度教育プログラムとして、限定メンバーに、毎年2回4ヶ月間金曜(応募者が多く短縮された)に、問題主導型の ESA 連続セミナーが開催されている。

さらに、5年以上の実績を持ち、NIH のリーダーになるという強い希望とポテンシャルを持つ人材を対象に、契約や助成のマネジメント・スタッフにも開かれた、NIH リーダーシップ開発指導(mentoring)というプログラムがある。参加者は他の研修プログラムを履修していることが必要である。

B-1.4 ピアレビュー・システムの見直し体制

NIH では常にピアレビューを含む評価システムを見直している。NIH の部局横断的な委員会である Peer Review Oversight Group (PROG) がすべてのピアレビューの調整・評価およびピアレビュー・ポリシーに関する勧告作成をしている。CSR センター長および CSR Advisory Committee (CSR のポリシー開発に多様な外部インプットを得るために設置)の委員長は PROG のメンバーであり、連絡調整にあたっている。

ここでの議論には、評価人材の育成・活用にあたって示唆的な内容が含まれている。

- パネルとレビューアの最適作業量
- レビューアの適性と資質・能力、多様性の向上
- パネル・マネジメント上のメンバー数制限
- 新人/非常勤レビューアに対するトレーニングのポイント

これらの評価調査、提言や議論の例を、B-1-資料2、B-1-資料3にまとめた。

◇ B-1-資料1 NIH グラント申請のピアレビュープロセス (R01 定期研究プロジェクトグラントおよび R21 開発的グラントの場合)

NIH のグラントのピアレビュー・システムは、第一段階の科学技術の質的側面に限定した評価プロセスをもつが、このプロセスを透明に公正に行うための様々な工夫が埋め込まれている。評価マネジメント組織に所属する評価人材は、このプロセスの的確な運用を担うことになるが、その具体的なイメージが様々な公開文書で明らかになっている。本事例 B1-資料1には、参考のためそれを取りまとめた。

出典：A Straightforward Description of What Happens to Your Research Project Grant Applications (R01/R21) After it is Received for Peer Review (<http://www.csr.nih.gov/REVIEW/peerrev.htm>)

出典：Review Procedures for Scientific Review Group Meetings (<http://www.csr.nih.gov/guidelines/proc.htm>)

(注)このイントロダクションは R01/R21 について説明したものであるが、ほかのタイプの CSR がレビューするグラント申請も同様の仕方で行われている。ただし、若干の相違点があり、数種の申請（たとえば、小ビジネスイノベーション研究 (SBIR: Small Business Innovation Research) およびフェローシップ) は迅速レビューを受けるので、R01 よりも 1,2 ヶ月受理締め切りが遅い。また、SBIR は常に特別重点パネルによってレビューされ、フェローシップはどのレベルでもパネルでの評価を見送られることはない。

▶ ピアレビュープロセス

1. 申請の受理・受付（毎年 10/11 月）

- 受理審査部 (DRR: Division of Receipt and Referral) が申請を受理。毎年 1000 通前後の申請が届く。

(主要なグラント申請の受付日には、配送トラックがグラント申請の入った数千の封筒を Rockledge 2 Building の集荷場で荷下ろしする。このビルは、NIH の CSR 科学レビューセンターの本拠である。封筒が開封されると、申請には受付日がスタンプされ、追跡できるよう NIH データベースに登録される。)

- 審査官 (Referral Officer) が、書面のガイドラインを使って申請の内容をチェックし、どの統合レビューグループ (IRG; Integrated Review Groups) がその申請のピアレビューを行うのにふさわしいかを審査する。また、審査官は申請が認められた場合、NIH のどの研究所/センターが資金提供するかもあわせて検討する。

(グラントサイクルごとに十数人の審査官 (Referral Officer) がレビューし、担当 IRG を決定する。IRG は、同種の科学分野をレビューする Study Section の集まりである。IRG が決定されたら、その IRG 内の構成要素である Study Section のひとつに申請は割り当てられる。IRG への割り当てに加えて、もしその申請に十分にメリットがあると認められた場合、どの NIH の研究所/センター (IC: Institutes/Centers) が、その申請に資金提供することがもっとも適しているかも審査官は決定する。IC が決定されたら、各申請には一意の申請番号が割り振られる。審査官は、Study Section と研究所への割り当てを行うために申請者からの書面のリクエスト (申請書に添付されるカバーレターさえも含めて) を真剣に考慮する。)

- その上で、IRG 内の Study Section の SRA をまじえて、どの Study Section がピアレビューを実際に行うかを検討する。

(この割り当てプロセスは関与者全員が同等の権限を有しており、必要な場合には、審査官と Study Section の SRA、研究所のプログラム代表者、申請者の間での相互調整がなされる。)

2. Study Section へのピアレビューの割り当て（毎年 12 月初め）

- ピアレビューを担当する Study Section と、資金提供する研究所/センターが決定

される。この決定結果は、コンピュータ合成の手紙として申請者に通知される。この通知内容に疑問がある場合、申請者は担当の審査官および SRA に連絡して質問ができる。

(申請の割当て完了後 10 日以内に、各申請者と資金提供を受ける研究機関に対して、Study Section と潜在的な資金提供研究所のリストを含むコンピュータが自動的に打ち出した手紙が郵送される。この通知を受け取ったら、申請者は、Study Section の SRA もしくは審査官 (301-435-0751) に連絡して、Study Section や I/C の割り当てについて質問できる。各 Study Section でレビューされる科学の内容や協会について定義する公式ガイドラインはあるものの、しかし、さまざまな Study Section によってレビューされる科学には重複もある。実際、今日研究プロジェクトのスコープが広がっているために、ひとつの申請がふたつ以上の Study Section に適している場合も多い。そこで、CSR スタッフは、申請をもっとも適切な委員会ひとつに絞って割り当てようと務めている。レビュー・ラウンドあたり 1000 通の申請すべての審査には最大 6 週間かかることがある。申請者が 6 週間後になっても何も通知を受け取らない場合、審査部 (Referral Office) に連絡したほうがよい。)

- SRA は、申請書の内容を読み、どのレビューアにレビューを依頼し、レビューの指定討論者に誰を当てるかを検討する。典型的な場合では、2、3 人のメンバーが各申請の書面レビューを割り当てられ、一人か二人のメンバーが討論者の役割を務める。必要な場合には、SRA が議長やレビューアと相談したうえで外部意見を求める。レビューに特別な専門知識・技能を要求する場合や、特別な状況 (たとえば、利害衝突が生じた場合など) のために、特別重点パネル (SEP: Special Emphasis Panel) がアド・ホックに設置されることがある。

(申請が Study Section に割り当てられたら、SRA は申請を読み込み始める。SRA はその内容を分析し、完成度をチェックし、どの Study Section メンバーが各申請をレビューしたり、討論者の職務を果たしたりするのに最適かを決定する。)

【注意】公式の CSR Study Section は、一般的に 18~20 人の個人から構成される。この個人は、SRA によって生物医学分野における活動的で生産的な研究者の中からノミネートされ、複数年にわたって役職を務める。この目標は、グループの複合知識が、レビューのために Study Section に割り当てられた研究対象・内容の多様性をカバーできるようにすることである。しかし、これは達成困難であるので、Study Section メンバーの機能は、非常勤メンバーや外部の書面での意見によって補完されることが多い。

3. レビュー資料の送付 (毎年 12 月末/1 月)

- レビューアと指定討論者が決まったら、すべてのレビュー資料が Study Section のメンバー全員に送付される。ただし、利害衝突があると考えられるメンバーには送付されない。

この送付は、Study Section 会合の 6 週間前に行われる。

(申請の提出からレビューまで数ヶ月にわたることから、申請者は補助的な資料を提出したいと思うことも多い。しかし、このような追加資料の受理について各 Study Section はポリシーを設けている (たとえば、長さや提出時期)。SRA に注意を喚起するためにも、受理可能な内容・形式・締め切りを確認するためにも、提出前に SRA に連絡を取ったほうがよい。)

4. レビュー対象の選定依頼（会合開催1週間前）

- Study Section の会合1週間前に、SRA は全メンバーに対して、科学的メリットという点から上位半分のランクに達しないと思われる R01 申請のリストを提出するように依頼している。上位だけがピアレビューにかかることになる。これらの下位半分の R01 申請は「見送り」(streamlined; 一定水準以下と評価されるものをレビューからはずし簡略化する) 状態にあるとみなされる。会合ではこれらの申請は採点されたり、議論されたりしない。レビューアの書面での批判が提供されるので、申請者はその後申請を修正して再提出してもよい。拒絶とは異なる。

(ある申請が上位半分ないと二人のレビューア/討論者が同意した場合、その申請は見送られる)

5. 最終「見送り」リストの送付（会合開催2,3日）

- SRA がメンバーから回収した「見送り」リストを統合し、最終リストをメンバーに送付する。

(このリストを閲覧した後に、利害関係のないレビューグループ・メンバーの誰かが同意せずに、彼/彼女が上位半分に属するので完全な議論を行うべきだと思う申請を決めることも出来る。)

6. Study Section 会合の開催（2月中頃～3月中頃）

- Study Section 会合開催。会合は、通常2日間にわたる。議長と SRA が共同で会合を運営する。NIH 内の研究所の代表者がオブザーバーとして参加する。レビューアと指定討論者が評価を提示し、その後外部意見がある場合は読み上げられる。一般的議論の後、メンバーは SRA が配布した採点表に各申請の優先順位得点 (priority scores) を個人ごとに書き込む。これらの表は、会合の最後に SRA もしくは管理アシスタントが集める。

(小さなバリエーションは多少あるものの、通常の CSR Study Section 会合はすべて同一の形式にしたがう。メンバーは相互作用が最大化されるよう会議テーブルを囲んで会合する。議長 (Study Section メンバーの一人) と SRA がともに座り、共同で会合を運営する責任を負う。NIH 研究所の代表者の席は会議テーブルの後ろ側であって、議論には参加してはならない。)

- レビューアに対しては留意すべき注意が与えられる (利害衝突の回避等の項目参照)。

<Study Section 会合の進行 (一般的パターン) >

- 1) 会合の冒頭、「見送り」とされた申請のリストが最終調整のため、大きな声で読み上げられる。
- 2) 「見送り」にあるとされた申請について、利害関係のないメンバーが得点に疑問を投げかけたり、コメントしたいと希望したりする場合には、通常のレビュー手続においてレビューグループ全体によって議論され、検討される。
- 3) 科学レビューグループの議長は、議論の対象となる各申請を紹介し、SRA が割り当てた各個人に対して、その評価を提示するよう依頼。
- 4) レビューアが意見を陳述する。
- 5) その後、指定討論者がコメントを求められ、グループの議論が続く。
- 6) 十分な議論を尽くしたうえで、議長がその申請に優先順位得点を割り振るよう求める。採点は、科学レビューグループの定期指定メンバーおよび常勤メンバーとして役務につく者によって行われる。
- 7) 採点シートは会合の最後に SRA が回収する。

<Study Section 会合での採点方法等>

1) 採点方法

－採点は、5つのレビュー基準（重要性とアプローチ、イノベーション、研究者、環境）を考慮したとき、その分野に対して与え得るインパクトの全体を反映するようになされる。

（なお、申請の性質やその相対的強みに応じて、ある申請と別の申請では異なる各基準に関する強調点がこの採点には付されている）。

－採点の最高点は100、最低点は500。

－個々のレビューアは2.2のような小数を含む2桁の数値で得点をつけ、個々のスコアを平均した上で、100をかけて採点された各申請の全体得点を求める。

－棄権したメンバーや議論の間に得点を提示しなかったメンバーの採点は、全体得点を求める際には用いない。

－見送りとされる申請が全体の半分を占める場合、残りの申請には100～300までの点が与えられる。また、もし申請の25%のみが採点されない場合、残りの申請は100～400の間の点が与えられる。（Study Sectionがどのくらいの申請を採点しないかによって、スコアの範囲は変わってくる。）

2) レビューアの棄権

－レビューアは棄権しないことが望ましい。しかし、議論の結果判断延期がふさわしいと考える場合、または、申請のメリットをアセスメント出来ないと感じられる場合、レビューアは投票シートに「AB」と書くよう勧められる。

3) 会合におけるレビュー見送り制の導入

－実質的議論の前に、科学レビューグループが議論される申請が実際には上位半分にはないと判断した場合、その申請は採点しないよう勧告される場合もある。そのような指示を行うためには科学レビューグループの全員一致の賛成が必要である。

4) 申請への倫理的配慮

－研究対象としての人間の利用や、動物の厚生、バイオハザードに関してコメントや重大な懸念がある場合には、これらのコメントや懸念を反映する規制をその申請に加える動議を導入でき、適切な指摘をサマリー・ステートメントに含めることができる。

5) 追加情報の要求などの発議

－レビューグループが勧告を作成する前に、追加情報が必要な場合、延期動議を発することができる。レビューグループは、多数の投票によって追加情報が得られるまで申請の審査を延期したり、申請を評価するのに必要な追加情報が施設の視察によってのみ得られたりする場合には、プロジェクト施設の訪問を発議できる。メンバーのうちの誰が延期の発議をしてもよい。

7. パーセンタイルの算出と通知（会合の数日後）

・すべての優先順位得点情報が申請データベースに入力され、パーセンタイルが算出される。このパーセンタイルは申請者に郵送される。

・パーセンタイルは、個別 Study Section の採点の中で優先順位得点が相対的にどれくらいか、どのランクにあるかを示す数字である。

（会合の数日後、申請データベースに入力されるとコンピュータが優先順位得点を算出し、パーセンタイルが申請者に自動的に郵送される。）

8. サマリー・ステートメントの作成（2月～4月）

・会合後、SRA がサマリー・ステートメントを作成する。Study Section ごとに平均80件のサマリー・ステートメントが作成される。サマリー・ステートメントは、レビューアの書面コメントと Study Section 会合でのメンバー間での議論について

のSRAのサマリーを統合したもの。Study Sectionの勧告や予算勧告、特別に考慮すべき条件についての管理官の意見を含む。サマリー・ステートメントの作成には約6週間かかる。

(申請者へのフィードバックは重要であるが、多数の案件で時間がかかる。)

9. サマリー・ステートメントの送付(4月末/5月)

- サマリー・ステートメントが、資金提供を行うにふさわしいとされたNIHの研究所に送付される。この時点で、Study SectionのSRAのコントロールは終了する。以後、レビューの解釈と申請の扱いに関する申請者とNIHの連絡役は該当研究所のプログラム・オフィシャル(Institute program official)となる。

(SRAの注意は、この後は次のグラント申請サイクルへと向くことになる。)

10. 研究所審議会による勧告の検討(5月/6月)

- 研究所審議会(Institute Advisory Councils)が会合を開催し、Study Sectionの勧告を検討する。

(研究所審議会はNIHのピアレビューの第二段階である。)

11. 資金受領(9月~11月)

- 研究所審議会のレビューを経て、資金授与を認められた申請者は資金を受け取る。

▶ 科学ピアレビューにおける公正さの確保—利害衝突の回避

科学レビューグループ会合に先立ち、各レビューアは利害衝突および秘密保持証明書と、レビューされる申請のリストを受け取る。会合よりも前にレビューアはSRAに対して、どのような利害衝突であっても通知し、レビュープロセスにおいて秘密が保持されることを確認しなければならない。

各会合の冒頭、SRAは、NIH利害衝突ポリシーを説明し、メンバーにオリエンテーションを行う。

レビューアは、申請を提出した営利組織のメンバーであったり、資金的利害をもっていたりする場合、申請の議論の間は退室しなければならない。利害衝突関係の見かけでも与える場合には、申請の議論の間退室しているべきである。

科学レビューグループ会合の終わりに、SRAはすべてのメンバーから、自身の出席が実質的もしくは見かけ上の利害衝突を構成する場合には申請のいかなるレビューにも参加しなかったこと、そして行動の秘密が保持されていることを示す証明書を集める。加えて、各Study Sectionは、どのアプリケーションに対して、どのメンバーが潜在的な利害衝突を理由に退室したかを示すログを保存する。このログはグラント・アシスタントが用意し、Study Section事務室が維持管理に当たる。

具体的な判断事項として、以下のような例がある。

—実質的な利害衝突がある場合

- 1) メンバーの所属機関が提出した申請が議論される場合。
- 2) メンバーもしくはその家族、職業上の仲間が申請内容に資金的利害や所有権上の利害をもつ場合。
- 3) メンバーが、申請の筆頭研究者の機関にサービスやセルライン、被験者、その他資料の提供者である場合。
- 4) メンバーが照会状の書き手である場合。
- 5) メンバーが申請を提出した営利組織のメンバーであったり、資金的利害をもつ

ていたりする場合。営利組織の株式を所有していたり、コンサルタントを務めていたりする場合も含む。

一見かけ上の利害衝突がある場合

- 1) メンバーの最近の学生や恩師、近しい個人的友人の申請が議論する場合。利害衝突がありえるかどうかは、公刊物などを頼りに、メンバーと筆頭研究者との間の関係の頻度や時間的近接をもとに判断する。
- 2) メンバーの研究室で現在進行中の仕事とほぼ重なるプロジェクトである場合。
- 3) プロジェクトで表明されている主張がメンバーと大きく異なっており、この事実がメンバーの客観性に影響があると合理的に見なされる場合。
- 4) レビューの組織や研究室に実質的な資金提供を与えている営利組織からの申請を議論する場合。

なお、多数の部局から成る学術機関や病院、保健所、研究機関の場合、独立したある部局所属のレビュー（連邦政府職員を除く）は、ほかに実質的・見かけ上の利害衝突がない限りは、別の部局からの申請をレビューすることができる。

また、ある個人が申請者に対して資源やサービスを提供しているものの、この資源やサービスが科学コミュニティの誰にでも自由に利用できる場合は、資源を提供する機関も個人も利害衝突にはない。

ほかの機関所属の参加個人を指名して申請が提出された場合も、この個人は利害衝突を構成する申請者機関と関係を持つとはみなされない。したがって、(1) 指名された個人は申請機関のほかの申請をレビューできるし、(2) 指名された個人の機関に所属するほかの個人も、いかなる実質的もしくは見かけ上の利害衝突があると考えられない《限りは》、提出された申請のレビューアとして活用できる。SRA は、利害衝突が存在しないという事実を文書化する。

▶ 科学ピアレビューにおける公正さの確保—秘密保持と研究者とのコミュニケーション

レビュー対象となる申請に直接関係のある資料すべては、コンサルタントと NIH スタッフのみの使用を目的とするためのコミュニケーションに限定されるものであって、ほかの個人に見せたり、ほかの個人と議論をしたりするべきではない。レビューグループ・メンバーは、直接関係ある IRG 外部の専門家に個別の申請もしくはその一部に関する意見やレビューを独立に求めてはならない。しかし、SRA が助言を得ることができる科学者をメンバーが示唆することはかまわない。コンサルタントは、レビュー会合の終わりに SRA のもとにすべてのレビュー資料を残して立ち去るよう要請される。グラント申請のみに限定された情報は、レビューアの利益のために用いたり、誰かと共有したりしてはならない。

どのような条件があつたとしても、コンサルタントは研究者やその組織、その他勧告対象となる誰であろうと助言を与えてはならないし、レビューの進行について議論してはならない。研究者は、未成熟であつたり誤っていたりする情報をベースにして未定見な行動を起こすかもしれない。同時に、そのような助言は進行手続の限定的性質への不公正な介入を示しており、レビュー委員会に参加するフェローコンサルタントや施設訪問チームのプライバシーを侵害している。秘匿性の侵害は、レビュー委員会への質の高いコンサルタントの参加を妨げ、委員会に参加した人々の勧告を目的とする自由で十全な議論への献身的参加を妨害する可能性がある。

施設訪問の間を除き、コンサルタントと研究者との間には直接のコミュニケーションが

あってはならない。コンサルタントの研究者に対する追加情報や電話問合せ、連絡の要請は SRA に対してなされなければならない。SRA がそのようなコミュニケーションのすべてを取り扱う。

以上をまとめると、次のように秘密の保持がなされるかたちがとられている。

- レビュー資料は、レビューアと NIH スタッフのコミュニケーションのみを目的としている。この資料をほかの個人に見せたり、ほかの個人と議論したりすることは禁じられる。
- 外部の意見を求める場合は、SRA を通じて行う。誰に助言を求めたらよいか SRA に示唆を与えることは許される。
- レビュー会合の終わりに、すべてのレビュー資料は SRA に返却される。
- 施設訪問時を除き、レビューアと申請者との間にはコミュニケーションがあってはならない。レビューアは、申請者やその所属組織に対して、個人的に助言を与えてはならない。また、レビューの進行状況についても議論してはならない。
- 申請者に対する追加情報や電話問合せ、連絡の要請は SRA に対して行う。

▶ 科学ピアレビューにおける公正さの確保—科学上の不正があった場合の処置

「不正」もしくは「科学上の不正」は、42 CFR 50.102 に次のように定義されている。データ捏造やデータの曲解、剽窃など、研究の提案・実行・報告について科学コミュニティ内部で共通に受容されている実践からの重要な逸脱であるその他の行為。これには、データの解釈や判断における正直な誤りや正直な差異は含まれない。

グラントや協力同意申請や契約プロポーザルの科学上のメリットについてのレビューは、通常保留や進行中の問合せ、調査によって遅延されることはない。レビュープロセスへの影響を避けるために、HHS 資金授与部門 (HHS awarding units) は、科学レビューグループ・メンバーに対して、ありえる不正や現在進行中の調査状況について情報を与えない。

しかし、確実な証拠がレビューの信用が毀損されるかもしれないほど明らかになってきた場合には、CSR 研究正当性保全官 (CSR RIO: CSR Research Integrity Officer) は、部局研究正当性保全連絡官 (ARILO: Agency Research Integrity Liaison Officer) と問題について話し合う。

科学レビューグループは、申請者の誰かの不正が明らかになった申請についてはレビューを中止する。不正の証拠が表面化した申請についてはレビューするべきではない。

調査が完了し得られた知見は、客観的で首尾一貫したレビューのために事実の正確な情報公開が必要である場合には、科学レビューグループにも共有される。SRA は CSR RIO に対してその証拠を報告すべきである。RIO は、適切な CSR スタッフと ARILO をその証拠をどのように扱うべきか決定するプロセスに関与させる。

疑わしい不正があるすべてのケースにおいて、SRA はレビューアたちにそのような証拠の深刻さと、秘匿性が厳重に保持されなかった場合に生じる潜在的な危害を協調すべきである。加えて、決定的となった疑いは HHS が重大なものとして取り上げ、追究することをレビューアに保証することが SRA は重要である。

どのような場合でも、SRA もしくはレビューアが科学レビューグループの懸念を筆頭研究者や申請機関に対して直接伝えてはならない。

◇ B-1-資料2 CSRによるIRGの評価改善活動(1)

1999年より本格的に開始されたCSR (Center for Scientific Review) のWorking GroupsによるIRG (Integrated Review Group) の評価・改善調査では、レビューアの負荷とレビューの質の関係、科学分野とIRGとの対応、レビューアのマネジメントなどについて評価と勧告が行われた。本報告では、この評価と勧告を中心に報告する。

出典：Peer Review Notes, Jan. 2002 掲載の評価・勧告の要約

▶ CSRの評価改善活動の経緯

CSRは、ピアレビュー主体であるIRGとstudy sectionの評価と改善を継続的に行っている。1996年以降では主要な活動は、1996年より開始されたWorking Groups of CSR Advisory Committeeと、1998年に開始されたPanel on Scientific Boundaries for Review (PSBR)によって担われている。

Working Groupは、1996年に細胞発達・機能IRGの評価を行うアド・ホックな組織として始まった。その後、1998年には、CSR Advisory Committeeの勧告により、すべてのIRGの評価・見直しが行われた。

▶ Working Groupsによる評価・勧告のハイライト

1. レビューアの作業負荷とレビューの質

- 一つのStudy Sectionの最適作業量は、レビューサイクルあたり60～80の申請の審査である。レビューの質やレビューアの士気をもっとも高い。最適作業量が上記に当てはまらないStudy Sectionは、分野の重複や不適切な分野区分が行われている可能性がある。
- レビューア一人あたりで見ると、10を超える筆記レビューを割り当てたり、筆記レビューと読解レビューの合計の割当てが14を超えたりする場合は負担が大きすぎる。

2. IRGの再編成

Work Groupsは、研究分野の重複や作業負荷のアンバランス、新興分野と斜陽分野のスクラップアンド・ビルドという視点から次のIRGの再編成を示唆した。

－研究分野の重複や作業負荷のアンバランスが見られるIRG

分子神経科学、細胞神経科学、発達神経科学、内分泌学・生殖、栄養学・代謝

－分割が必要なIRG

臨床研究、筋生物学研究、脂質／リポ蛋白研究

－新興分野に対応し新設すべきIRG

インフォマティクス研究、プロテオミクス研究、ゲノム研究

• CSRの対応

IRGの再編成作業は、前出のPSBRが継続して行っている。また、すでに神経科学IRGの再編が行われたほか、腫瘍科学IRGに病理学C study sectionが創出されるなどの調整がすでに行われている。

3. レビューアの適性と資質

- レビューアの専門知識・技能や適性、公正さは一部の例外を除き卓越している。しかし、より多くのシニア・レビューアや臨床家、女性、マイノリティーが参加することが望ましい。
- レビューアに対して、よりフレキシブルな業務条件や資金的支援の拡張などより高いインセンティブを与えるべきである。

4. レビューアのトレーニング

- 新人レビューア／経験不足のレビューアに対して、レビューのトレーニング・ガイドラインが必要である。
- CSR の対応
現在、レビューアのトレーニング手続きを開発中。

5. パネルのマネジメント

- Study Section のパネルのレビューアが 30 人を越えると、会議の進行に問題が生じる。
- RA とパネル・チェアは新人／非常勤レビューアに対して、次の指導を行うべきである。
 - 1) 批判の準備を行うこと
 - 2) 方法論的詳細へのフォーカスを強調しすぎないこと
 - 3) 申請の採点の仕方
 - 4) パネルの間に効率的に申請を批判すること。
- パネル・チェアに対してピアレビューの実践とポリシーと同様、グループ・ダイナミクスとコンセンサス形成について追加的トレーニングを提供すべきである。

▶ 勧告と CSR の対応からの結論

ピアレビュー・パネルのマネジメントに関しては、次のことがこの勧告から学ぶことができる。

- パネルとレビューアの最適作業量
- レビューアの質と多様性の向上
- パネル・マネジメント上のメンバー数制限
- 新人／非常勤レビューアに対するトレーニングのポイント

また、2000 年～2001 年にかけて、SRA とパネル・チェアのベスト・プラクティス・ガイドラインがまとめられている。これは SRA とパネル・チェアが、レビュープロセスの各段階で何にどのように留意すればよいか述べたものである。これは、評価プロセス作成において参考になろう。

しかし、現時点では、CSR においても、ピアレビュー・レビューアのトレーニングについては確立されていない。CSR が調査・検討中のパネル・チェアとレビューアのトレーニングについては、継続して調査が必要と考えられる。

◆ B-1-資料3 CSRによるIRGの評価改善活動(2)

1998年9月CSR Advisory Committeeは、CSRが、IRGとそのStudy Sectionの組織とマネジメント、リーダーシップを評価するWorking Groupを設置するよう勧告した。IRGによってレビューが行われる関係分野の活動的で幅広い尊敬を集めている研究者を捜し求めてWorking Groupを組織した。2年間にわたりWorking Groupsは19のIRGすべてを評価し、CSR Advisory committeeに報告書を提出した。このレポートとCSRの対応について、以下に紹介する。

出典: Peer Review Notes, Jan. 2002

▶ レビュー対象となる科学のスコープと幅

Working Groupは、「作業負荷がstudy sectionの効果に対してインパクトを持つこと」を認めた。レビューサイクルあたり60~80の申請をレビューするstudy sectionは、最適レベルの責任とスコープをもつ傾向がある。Study sectionが90を超えてレビューをするか、もしくは50未満の申請しかレビューを行わない場合、レビューの質とレビューアの士気が悪影響を受ける。しかし、Working Groupが注記するところによれば、望ましい作業負荷であっても悪影響を受けるStudy Sectionもある。なぜならば、それらの「Study Sectionは分割されており、2つもしくは3つの小さなStudy Sectionのように運営されているか、あまりにも広い科学分野をカバーしているか、逆にあまりにもせまい科学分野しかカバーしていない」からである。

Working Groupは次のことを示唆している。少なくとも5つのIRGがカバーする分野が重複していないか、作業負荷のバランスはどうか、整合性はどうかについて検討すべきである。5つのIRとは、3つの神経科学IRGと、内分泌学・生殖IRG、栄養学・代謝IRGである。

Working Groupsは、新興の科学分野や斜陽の科学分野、分割の必要な科学分野にも注目した。同グループの指摘では、臨床研究および、筋生物学研究、脂質/リポ蛋白研究は分割する必要がある。インフォマティクス研究および、プロテオミクス研究、ゲノム研究は、CSRが取り扱うべき新興研究分野として認識された。

多くのWorking Groupの勧告と関心はCSR Study Section Boundary Teamによって採用され、IRGの再編成に取り掛かっている。しかし、CSRはすでに複数の新しいStudy Sectionを創出し、現存のsectionの科学の境界を調整することによって、いくつかの勧告にはすでに対応している。病理学C Study Sectionが腫瘍科学IRGに付加され、新しい筋生物学Study Sectionが創出された。分子神経科学、細胞神経科学、発達神経科学各Study Sectionの境界も調整され、追加的な調整も多くのほかのIRGにおいて行われている。

▶ レビューアの適性と能力

少数の例外を除き、Working Groupsは、「レビューアの専門知識・技能や適性、公正さは卓越していること」を見出した。しかしなお、多くのStudy Sectionでは「シニア・レビューアや臨床家、女性、マイノリティーがより多くの参加することによって利益を得るだろう」と指摘している。また、「CSRはレビューアに対して、現在の資金的支援の拡張やよりフレキシブルな業務条件といったよりよいインセンティブを与えること」を勧奨

された。

資金支援期間の修正は NIH の認可が必要であり、NIH 所長は、コストやピアレビューされた研究のみに資金支援を与えるという原則を理由に、この提案を受け入れることに難色を示してきた。しかし、CSR は現在の NIH 所長とその Committee Management Office とともに、現行の役職構造を修正すべく作業を進めている。提案には次のようなものも含まれる。「よりフレキシブルな業務条件の確立、シニア・レビューア向けに新しい限定的な業務カテゴリーの創設、CSR がその多様性ニーズの取り扱いを支援する新しい人材募集戦略の確立。」

▶ パネル (review meeting) のポリシーと手続き、マネジメント

Working Groups は、CSR のオリエンテーション資料とプレ・パネル・プレゼンテーションの有用性を認めた。しかし、パネル・チェアと SRA は「新しいレビューアや非常勤レビューアに対して次の点で追加的トレーニングを提供すべきである」と示唆している。「1) 批判の準備を行うこと、2) 方法論的詳細へのフォーカスを強調しすぎないこと、3) 申請の採点、4) パネルの間に効率的に申請を批判すること」。また、CSR はパネル・チェアに対してピアレビューの実践とポリシーと同様、「グループ・ダイナミクスとコンセンサス形成について追加的トレーニングを提供する方法を調査すべき」と、同グループは示唆している。

Working Groups は、「10 を超える筆記レビューを割り当てたり、筆記レビューと読解レビューの合計の割当てが 14 を超えたりする場合は負担が大きすぎることを」を発見した。Study Section の「パネルが 30 人を越えるレビューアが参加するときはやはり問題である」。

申請の、「スコアをつけないレビュー」のポリシーやプロセスについてはさまざまな意見があり、Working Groups の中には CSR はこの実践を修正すべきだと勧奨するグループもあった。しかし、Working Groups は「モジュラー予算研究グラント申請については承服できない」という点で全会一致した。とくに、予算の正当化の不足は問題である。CSR は NIH Office of Extramural Research とモジュラー予算研究グラント申請について議論を行い、早急にこのポリシーを評価するよう助言された。

(注) 「モジュラー予算研究グラント申請」の目標は、次の2つとされている。

- 1) 研究プロジェクトグラントを支援メカニズムとして再定義する
- 2) 申請者とピアレビューアを複雑な予算の作成・評価・交渉の義務から解放し、彼らを提案された研究の科学的・技術的メリットに集中させる。

このグラントは次のような特徴をもっている。

- 1 年あたり 25 万ドルまでのグラント予算について適用される。
- 研究の適切な支援を反映する上限 2 万 5 千ドルのモジュールに直接コストを分割する。この予算は後年度増加することはない。
- 一年によってモジュールの数は増減しない。
- キーパーソンについて、そのレベルと役割、貢献について予算上の説明を行う、など。

このグラント予算の詳細は、次を参照されたい。

(<http://grants.nih.gov/grants/funding/modular/modular.htm>)

CSR は、SRA 向けのベスト・プラクティス文書を開発してきた。この文書には、レビューアの割当てのガイダンスを含むパネル・マネジメントの幅広い助言が含まれている。2001 年 12 月、CSR は「新人レビューアと非常勤レビューアの訓練についての SRA ガイド」の開発を開始した。また、CSR は Study Section チェアのトレーニング選択肢も調査中である。

申請の“スコアを使わないレビュー”についての関心を扱うため、CSR は次のことを考慮している。1) 電子的なレビュー・モジュールにより依存する、2) これらの申請の弱点に関する簡潔な Study Section での議論を行う、3) パネルの冒頭にこれら申請に関してレビューア同士がどう考えているか情報共有を行えるようにする。

▶ **新しい研究方向と新興分野の調停**

Working Groups は、「多くの申請が革新的ではあるものの、新規性やハイリスク・ハイインパクトと考えられるものは少ない」と記述している。Working Groups のメンバーは、これらの申請は適切にレビューされていると一般的には感じているものの、「新人レビューアや経験不足のレビューアは、方法論的関心に不必要に拘泥する傾向がある」とも指摘している。すでに指摘したように、CSR は効果的批判という要素も含む新人レビューアのトレーニング手続きを開発中である。

(Working Groups のコメントのより詳細なサマリーは、CSR のウェブサイトです入手可能である。<http://www.csr.nih.gov/NewsFlash/newsflash.htm>)

事例 B-2. 欧州の行政機関等の評価とその人材育成の取り組み

欧州の行政機関・資金配分機関等での評価活動の概要と、それらに在籍している評価人材の育成の状況について、事例的に以下にまとめた。必ずしも評価専門人材を確保している機関ばかりではない。

敢えて特徴づければ、評価の事務局機能が主たる場合には、育成プログラムをもっておらず、評価の専門性を機関内外から調達している傾向が、また、評価機能をもつ場合には、評価の狭い専門性よりは、博士号保持などの研究開発上の基盤経験・能力の保有者を採用したり内部募集する傾向がみられる。OJT が意識され制度化されているケースもある。実務的なレベルでの訓練は別にして、Twente 大学などへの派遣研修は評価されている。PREST でも派遣・研修経験事例がみられる。

B-2.1 オランダ 技術財団

(STW:Technology Foundation STW)

1. 機関概要

- ・オランダの大学や企業の研究者に対する研究開発費の交付機関である。1981年に設立された。
- ・年間予算は約4,500万ドルである。
- ・Ministry of Economic Affairs と Ministry of Science and Education の双方から予算を得ている。8割が教育科学省のプロジェクト、2割が経済省のプログラム関連である。
- ・新規プロジェクトは毎年約100件あり、約250件の応募の中から事前評価により選択している。

2. 評価について

- ・実施している評価は、事前評価、モニタリング、中間評価、事後評価、追跡評価の5種からなる。

(1) 事前評価

- ・それぞれのプロジェクトごとに、4人以上7人以内の評価者によるピアレビュー（評価委員会）を行っている。
- ・評価委員会は、プロジェクトごとに異なる評価者構成となっている。
- ・一つの委員会は、STWからの事務局員と当該分野の専門家及び当該技術を利用する可能性のあるユーザーから成っている。
- ・メールレビューも実施している。

(2) モニタリング

- ・年2回程度、各プロジェクトの進捗状況の確認を行っている。

(3) 中間評価

- ・プロジェクト期間はそれぞれ異なるが、平均すると約4年である。

- ・ 2年経過した時点で、事前評価と同様の構成（但し、評価委員会はパネルレビュー）の評価委員会を組織し、今後とも研究開発を継続するかどうかの評価を実施する。
- ・ 委員会は、知的財産権等の関係から原則として非公開としている。

(4) 事後評価・追跡評価

- ・ プロジェクトが開始してから5年後と10年後に実施する。
- ・ STWの事務局員が、当時研究に関わった研究者や利用者から、電話による聞き取り調査を実施する。
- ・ 定量的評価については実施していない。

3. その他

(1) 評価結果の予算等への反映方法

- ・ 中間評価は厳格であり、これで、研究の継続・中止が決定される。
- ・ 事後・追跡評価は、国民への情報開示のために行われる。
- ・ STWの予算は、STWが行う個々のプロジェクトやプログラムの評価結果と直接的には関係はない。
- ・ 10年前に政府によりSTWの機関評価が行われ、良い評価を得た。
- ・ 予算は増加しているが、これは科学技術の重要性と2つの省からSTWが信頼を得ている事が理由となっているようである。

(2) 社会経済性インパクトの評価

- ・ 定量的アウトカムの測定については、Sussex大学のSPRU (Science Program Research Unit) で調査した結果があるが、研究自体とアウトカムとの間関係は間接的であり、測定困難であると考えられる。
- ・ STWでは、知的財産権の状況や定性的な評価をヒアリングの実施により得ている。

(3) 参加企業の評価

- ・ 企業自体の評価はしていない。

(4) STWの競合相手

- ・ STWは、NWO (Netherlands Organization for Scientific Research (オランダ科学評議会)) という政府組織に属しているが、どちらかというとなWOは教育科学省との関係が親密である。資金は、教育科学省からNWOを経てSTWに入る。オランダ経済省は直接に資金をSTWに支出している。
- ・ オランダ経済省は"SENER"という組織を独自に持っており、それは特定のいくつかの提案公募型プログラム（経済省関連）を扱って、産業界および大学へ資金提供を行っている。この点において、"SENER"とSTWは競合関係にある。

(5) 評価人材育成

- ・ OJTの形で行っているといえる。Arie RipがTwente大学で行っているものは、科学研究に対する一般的なレベルの評価者能力を身につけさせるものだが、STWでは教育プロジェクトも評価する必要があり、より実践的なレベルでのトレーニングが必要である。
- ・ 過去に外部研究機関へ人を派遣したことはない。
- ・ 採用では、自然科学系人材の採用を行うことにしている。
- ・ OJT制度では、各セクションにcoachと呼ばれる指導者がいて指導を行っている。

B-2.2 ドイツ 連邦教育研究省

(BMBF ; Bundesministerium für Bildung und Forschung)

1. 機関概要

- ・ドイツ政府のR&D予算は168億マルクであり、BMBFの予算は、2000年において、その64.3%、108億マルクである。
- ・BMBFの予算のうち、約30%がマックス・プランク協会(MPG)やフラウンホーファー協会(FhG)等の研究機関への助成金であり、約59%がプロジェクトへの助成となっている。プロジェクトへの助成金が増加傾向にある。
- ・IT、バイオ等15のプログラムの下、約1万のプロジェクトが実施されている。

2. プログラムおよびプロジェクト評価について

実施しているプログラム/プロジェクト評価は、事前(ex-ante)、中間(accompanying)、事後(ex-post)、混合(mixture)の4種。

(1) 中間評価

- ・実施時期は確定していない。米国、日本等海外の情勢変化への対応等の見極めを行っている。

(2) 事後評価

- ・通常プログラムごと、5年ごとに実施し、期間は3ヶ月～1年半であり、プログラムにより様々である。個々のプログラムには20～50のプロジェクトが含まれる。
- ・Arthur D. Little、マッケンジー等の外部シンクタンクへの委託により実施している。評価のガイドラインは内部のものであるのでシンクタンク等外部へは知らせない。シンクタンク独自の知識により、個々のプロジェクトにおいてプロポーザルに示されたベンチマークについて評価してもらおう一方、研究開発内容、方法、コスト等について分析してもらう。
- ・なお、シンクタンクから提出される報告の一般的様式は存在しない。この報告は参考情報としてのみ用いられる。また、報告は結果が提出されるだけでなく、幅広い情報提供や議論など多様である。

(3) 混合評価

- ・中間、事後評価を個々のプログラムに応じて適宜組み合わせて実施。

3. 研究所の評価について

(1) 対象

MPG、DFG、FhG、Blaue Liste Einrichtungen、Helmholtz Association等。

(2) 評価委員会

- ・常設委員会において各研究所のシステムを評価する。委員の任期は5年で、科学者等に構成されている。
- ・2000年より実施する。2000年はブルーリストの83研究所の評価を行った。

(3) 基本的スタンス

- ・研究所が実施する個々のプロジェクトは実施機関の責任において実施するものである。
- ・BMBFの評価スタッフは少人数であり、現状では研究所が実施する個々のプロジェクト

ト評価までできる体制にはない。

(4) 背景

- ・この研究所の評価はドイツ統一後10年が経過し、より競争力を高めるため「東」の機関を評価することとしたが、それでは unfair であるので「西」でも同様に実施する事とした経緯がある。

4. その他

(1) 予算配分の決定について

- ・予算の大枠については、カウンスル (Science Council) から BMBF に調査結果が報告され、意見交換後に決定される。但し、カウンスルはプロジェクトレベルまでコントロールできるわけではない。
- ・BMBF は、実施者側との議論・調整を行い、その結果として、個々のプロジェクトの予算が決定される。
- ・研究テーマの内容は、BMBF が考える。
- ・たとえば、FhG であれば、90% が連邦政府 (BMBF) から助成され、残り 10% は地方政府等からの助成金である。

(2) プロジェクト等に参加した民間企業の評価について

- ・評価対象外である。

(3) 評価人材養成について

- ・評価に関係した研究制度は特に存在しないが、OJT の制度はある。
- ・PREST 等外部の研修期間へ行った人に対し基本的に費用は負担する。しかしながら、過去に行った人は実際にはあまりいない。組織としての派遣制度はない。

B-2.3 英貿易産業省

(DTI ; Department of Trade and Industry, Assessment Unit)

1. DTI 評価 Unit の概要

- ・評価 Unit は DTI において評価のみを行う唯一の部署である。
- ・評価 Unit の人数は 5 人であり、彼らは評価者でもある。
- ・LINK プログラムは全体としては 1000 のプロジェクトから構成され、また一つの LINK プログラムは 20~40 のプロジェクトから成っている。

2. 評価について

- ・評価の種類は、事前評価、モニタリング、事後評価の 3 種である。

(1) 事前評価

- ・新プログラムの推進を評価検討する。これはアプレーザル (appraisal) と呼ばれている。
- ・提案された活動の種類で、助成するプロジェクトをどう選定するか、また、DTI はプログラム全体の運営を任せる特定 contractor をどのように決定するのかについて、ROAME ステートメントの事前評価のセクションにおいて説明されている。

- ・提案されたプログラムが Innovation Program Committee および DTI の大臣により承認されると、プログラムは実行に移される。

(2) モニタリング

- ・現在進行中のプログラムの進捗状況に関する情報の収集・分析を行う。
- ・目的（おもにオペレーションレベルで設定された目的）に対し、プログラムの進捗が program manager により監視される。program manager は任期が 3~4 年で人が交代する。
- ・program 開始に必要なことは、program manager が ROAME ステートメントを提言し、この提言が承認されることである。

(3) 事後評価

- ・終了したプログラムの成果 (outcomes) を事後評価する。
- ・事後評価はプログラム終了時に実施する。
- ・プログラムの評価に使用する方法はプログラムの性質により異なる。LINK プログラムのような共同研究プログラムにおける評価では下記項目が含まれるであろう。
 - ①企業や大学が基金に応募するに至った経緯およびプログラムに関するビジネスプランのドラフトが ROAME ステートメントから作られる。
 - ②効果性、タイミングの良さ、公正性にもとづきどのように提案は処理され事前評価されたのかを検討する。
 - ③ROAME ステートメントの中に書かれているもともとの意図に最終的に選択されたプロジェクトがどれくらい対応しているのかを比較する。
- ・評価レポートは結論および提言 (recommendation) を含む。このレポートは将来のプログラムに対し政策的にフィード・バックされる。フィード・バックは DTI senior management の部署において実施される。

3. その他

- ・DTI の研究開発プログラムの資金は Innovation Budget から出ている。
- ・Strategic Evaluation Committee が評価 Unit より上のレベルでどのような評価が行われるべきか等に関しコントロールを行っている。しかし、この Committee は評価 Unit の助言に高度に依存している。
- ・今回面談した人達は過去に DTI 付属の研究所で働いていた人たちであった。彼らでは掌握できないような領域については経済学者や統計家などのスペシャリストの手助けがある。
- ・EMG(Evaluation Methodology Group)は評価に関する peer group であり、主に経済学者からなっている。EMG は outside contractors を使用する傾向にある。
- ・プログラムを分野ごとに一つの全体としてみる傾向にある。個々のプログラムの評価からは政策的な提言(recommendation)は生まれてこない。例えば、バイオテクノロジーの分野への DTI の資金提供がはたして適切なものかどうかを判断しようとしている。一つの分野におけるメカニズムを分析する。これに対して、program manager はプロジェクト評価を行う。もしうまくいかないプロジェクトが存在する場合には、ケーススタディを実施するし、成功したプロジェクトがあれば、その成功した原因を探ろうとする。
- ・DTI のプログラムが他省のプログラムと相まって、うまくすべての分野を抜けなくカバーできているかどうか配慮することが重要である。
- ・DTI の役割は産業支援である。
- ・あるプログラムが DTI によって実行された場合と、DTI 以外のものによって実行された

場合を比較し、結果の違いを評価することは困難である。

- ・プログラム・マネージャーによる link completion report において、民間企業の評価を書くことがある。民間企業の参加者を評価しようとしている。
- ・LINK の目的は知識と basic technology の創出であり、precompetitive な性格を有している。
- ・以前は IPC(Individual Programme Committee)が資金提供に関する決定の責任を担っていたが、もはや存在しない。
- ・各プロジェクトは以下のものにより評価される。
 - 1.criteria
 - 2.program manager
 3. program management committee (産業界、アカデミア、当該技術のユーザー)
- ・各プロジェクトに関し目的は以下の階層構造をなす。
 - 1.DTI は何をなすべきか。
 - 2.プログラムの戦略的目的
 - 3.プロジェクトの個別目的

(人材養成について)

- ・内部研修制度については特にない。統計学などの研修制度は存在する。
- ・過去に PREST に人を組織として派遣したことがある。
- ・評価人材を外部から採用することはない。内部募集によっている。

B-2.4 英 工学物理科学研究会議 (EPSRC)

1. 機関概要

- ・EPSRC のミッションは物理、工学分野の研究開発を支援すると同時にポストドクの教育を行うこと。英貿易産業省 (DTI) の office of science and technology の下に7つのリサーチ・カウンシルがあり、そのうちの一つ。
- ・年間予算 4 億 2000 万ポンド
グラント 3 億 2000 万ポンド
ポストドク教育 7000 万ポンド
その他 3000 万ポンド
- ・プロジェクトの平均期間は3年であり、平均金額は 150000 ポンドで、この金額は契約研究者の人件費を含むが、終身雇用の大学研究者の人件費は含んでいない。これと建物の費用は Department of Education からでている。
- ・11 のプログラムが存在する。そのうちの9つは EPSRC の program (化学：40 百万ポンド、物理：38 百万ポンド、数学：8 百万ポンド、IT：50 百万ポンド、材料：50 百万ポンド、工学：50 百万ポンド、製造 (機械)：25 百万ポンド、社会基盤：25 百万ポンド、ライフサイエンス：9 百万ポンド) となり、あとの残りは Basic Technology Program (2 年間で 41 百万ポンド) 及び E-science Program (3 年間で 190 百万ポンド) を他のカウンシルの代理で実施している。

2. 評価について

- ・行っている評価の種類としては、事前評価、事後評価の2種類である。

(1) 事前評価

- ・年間に約 1500 のピア・レビューがありその約 1/3 を採用している。下記の2方法によって審査を実施している。
 - ①郵便等によるピア・レビュー
 - ②パネルでピア・レビューの結果を用いながらピア・レビューの優先順位を付ける。
- ・100 のプロジェクトの審査に 10~12 人のパネラーがあたる。各パネラーは1年に 10 のプロポザール審査を行う。そのうち 1/3 に資金を提供する。

(2) 事後評価

- ・下記4つの評価基準を用いてピア・レビューを実施する。メールおよびパネルを使用する。産学の人達に評価を実施してもらう。
 - ①世界的レベルに比較して科学的質はどうか
 - ②社会的なニーズと比較して従事している研究者数ははたして適正なのか。従事者の技能は適正なものか。これらに関して産業界の人達に意見を求める。
 - ③研究インパクト：当該研究分野への貢献についてはどうか
 - ④商業化についての可能性

3. その他

- ・EU の法律により、民間企業への直接的な、国家資金提供は禁じられている。
- ・評価者は評価プロジェクトの金額変更を行えとは言わない。金額決定をするのは EPSRC である。
- ・1 プロジェクトの評価パネルにおいて評価者の数は3人である。50%は大学から、1/3は産業界からの人で構成されている。連続性を重んじ、プロジェクトの開始時と終了時で同一の人が評価するよう努力している。パネルは ad hoc であり、パネル会合時には、各評価者に1日 120 ポンド支給される。
- ・EPSRC はプロジェクト終了時に socio evaluation を実施しているが、追跡評価は実施していない。
- ・EPSRC が援助したすべての PhD Students をサーベイし、後に、援助した分野の仕事に就く事ができたのかどうかの調査が実施された。
- ・2年ごとに9つのプログラム分野の評価を行っている。
- ・評価レポートを見るのは2つの advisory councils で大学からの technology opportunity panel と産業界からのユーザーパネルに分けられる。
- ・2年ごとに6月に Treasury Comprehensive Spending Review が出される。これは、EPSRC の次期3年間の予算を定めるものである。例えば、「2002年 Review」により、2003年から2005年分が出されるが、このうち、2003年の分と2004年の分は固定されているが、2005年の分は unfixed である。次のレビューは2年後に「2004年 Review」として出され、ここで、2005年の分と2006年の分が固定され、2007年分はまだ unfixed されている。

(人材養成について)

- ・ある種の OJT 制度がある。約 45 人が評価の仕事をしているが、仕事内容の 15% が評価の仕事となっている。

- ・Twente 大学等に人を組織として派遣した。評価の人材養成プログラムとしては Twente 大学が最も良いであろう。
- ・職員を選ぶ時には評価業務とは無関係に選んでいる。採用に関しては（主に工学や物理科学の分野の）博士号を所持していることが必要となる。採用過程で、志願者に厚いレポートを送付し、その要約を提出させて審査の一助とすることがある。

B-2.5 フランス国立科学研究センター

(CNRS ; Centre National de la Recherche Scientifique)

1. 機関概要

- ・仏最大の研究機関であり、全科学技術分野を対象とした基礎研究を担当している。
- ・約 2.5 億 euro/年の予算を主として、Ministry of Research から得ている。
- ・職員数は約 25,000 人、そのうち約 11,000 が研究者である。
- ・CNRS 所属の研究所は約 150 あり、そのうち、関連のある（資金を提供している）研究機関は 1,300 である。
- ・各研究機関は CNRS 以外からも通常は予算を得ており、研究者も CNRS 所属と大学など他機関所属の研究者が混在している状況であり、緩やかな関係となっている。

2. CNRS における評価について

(1) 評価の概要

- ・主として機関評価の観点から評価を行い、1,300 の研究所全てが評価対象である。
- ・プロジェクト、プログラム、研究者についての評価も実施されている。
- ・プロジェクトの期間は 4 年間が基本となっており、事後評価を受けて継続の可否が決定される。
- ・CNRS においては 1939 年に評価が開始された。

(2) 委員会組織

- ・国家科学研究委員会（以下、委員会 = CNRS ; Comité national de la Recherche Scientifique）で評価を行っている。
- ・委員会は科学技術分野ごとに 40 のセクションに分割され、各セクションは 21 名の委員で構成されている。
- ・委員の内 14 名は研究者の中から幅広く選挙で選ばれ、7 名は大臣が指名する。
- ・外国人（EU が中心）や、産業界からも委員が選ばれている。
- ・CNRS が委員会の事務局となっている。

(3) 評価方法

- ・ピアレビューが基本となっている。
- ・CNRS のミッションは新たな知の獲得である。そのため研究の質の評価が主となっている。
- ・委員会は年に 2 回、4 日間かけて開催され、研究所訪問などを通じて評価を行っている。
- ・セクションごとに担当する研究機関について、4 年に 1 回評価を実施する（1 年に 1/4 ずつ評価していく）。

- ・評価の対象には研究所の運営、産業界との連携なども含まれる。
- ・委員会は研究者の評価、研究者採用に関するアドバイスも行う。
- ・評価により研究所をランク付けしている。しかし、あくまでもこれは評価及びアドバイスの位置付けであり意思決定とは区別されている。

(4) その他、質疑応答など

- ・評価の反映は難しい。例えば、担当者が替わらないと反映されない場合や、自ら進んで変革に活かす場合など様々である。
- ・評価計画の策定においては、評価目的、クライアントは誰か、を明確にすることが肝要となる。
- ・企業の個別評価に関しては、情報が企業内に存在するため難しい。
- ・CNRS の機能は評価委員会事務局である。それゆえ、評価専門家の採用や評価者としての育成、研修などは行っていない。
- ・グランゼコールは CNRS によって評価される事を気にしない。しかし、他の機関は CNRS により良い評価を得てそれをアピールの材料にしようとしている。

(5) 人材育成

- ・OJT 制度や研修制度等は存在しない。
- ・人を派遣する制度は存在しない。
- ・職員の選抜基準は評価とは無関係のものである。

B-2.6 フランス原子力庁

(CEA ; Commissariat a l'Energie Atomique)

1. 機関概要

- ・原子力とおよびその利用（エネルギー、産業、健康、防衛）の研究を行う公的機関である。
- ・平和利用目的の原子力エネルギーの競争力や、社会による受容の確保を目指した研究の他、エネルギー、IT、バイオ分野における研究開発も実施中である。
- ・主に、Ministry of Research から年間 1860,000,000 ユーロの予算を得ている。
- ・職員数約 16,000 人であり、そのうち 46%がマネージャー、技術者、研究者からなる。
- ・1,600 件の有効な特許を所有し、毎年 200 件ほどの新規特許を取得している。
- ・約 90 の新規事業が既にスピナウトして起業されていること等、原子力以外でも産業界へ貢献していることをアピールしている。
- ・4年（2001～2004）の契約を政府と締結しているが、この契約中に各分野のマイルストーン、成果計測指標などが明記されている。
- ・CEA は、下記の 4 セクターに分割されている
 - －原子力エネルギーセクター
 - －防衛セクター
 - －技術研究セクター
 - －基盤研究セクター

2. CEA における評価について

(1) 評価の概要

- ・評価の考え方は、下記2種類に分類される。
 - 1)科学的、研究的観点からの評価
 - 2)目標達成度に関する評価
- ・1)は委員会形式(カウンスル)による外部評価で行い、2)はマネージメントの観点から内部評価により行う。

(2) 委員会組織

- ・科学的観点から評価を行うカウンスルは、材料科学(10カウンスル)、バイオ科学(7カウンスル)、原子力(5カウンスル)、科学技術(1997~2000年は4カウンスル、2001~2004年は6カウンスル)の4分野に分かれ、合計で約300人の研究者(任期は4年間)が任命されており、2年毎に評価を実施する。
- ・委員構成は、CEA内部;3%、仏国内;66%、海外;31%となっている。
- ・さらに細かく32会議に分かれており、各会議は10~15名程度で構成されている。
- ・事前に資料を送った後で、会議をCEAで3日間開催してレポートを作成する(2回/年)。
- ・CEAでは基本的には成果や運営方法に対する事後評価を実施し、アドバイス、リコメンデーションなどを取りまとめる。
- ・同じ会議で重要なプログラムの事前評価を行う場合もある。

(3) 政府との契約

- ・上述の政府との契約が事前評価に相当し、4年ごとに締結するが毎年見直しを実施している。
- ・Ministry of Research, Ministry of Economy, Finance and Industry、CEAの間で契約は政府との数ヶ月に及ぶ交渉を経て、締結される
- ・契約においては研究内容とともに、研究の進捗に関するマイルストーンと成果計測指標(performance indicator)を定めるものである。
- ・この契約は事前評価であると同時に、モニタリング・ツールでもある。

(4) モニタリング(中間、事後評価)

- ・科学・研究面のモニタリング評価は上述のカウンスルにおいて実施される。
- ・マネージメントの一環として目標達成面のモニタリングは実施され、各セクションが3ヶ月ごとにレポートを作成し、マイルストーンに関する達成度、計測指標のデータなどを報告を行う。
- ・CEA全体としては毎年レポート作成を行い、政府へ報告している。

(5) その他、質疑応答など

- ・研究者は上司により、毎年評価されている。
- ・経済性の評価に関しては、特許のライセンス料、スピノフした研究者の数、産業界からの投資額などが技術移転の計測指標として政府との契約で定められ、これにより評価が実施されている。
- ・評価の情報は、CEA内部のデータベースや、研究関連の情報を収集する独立機関(OST)により提供されている。
- ・内部に評価専門家の育成制度は存在しない。外部派遣制度は存在しない。
- ・欧州南部においては評価専門家の伝統は存在しない。しかし、例えば米国のコンサルタント会社に評価に関し依頼するというような事はある。

事例 B-3 EU における科学技術ピアの公募・選定・行動規則等

プログラムに基づいて実施されるプロジェクトの評価では、公平性や中立性を維持しつつ意思決定の前段の評価作業を専門的見地から外部専門家（個人）に担ってもらうことが一般的である。その透明で厳正な手続きを文書化した代表的なものとして、EU の第 5 次フレームワークプログラム／RTD プログラムで用いられているマニュアルがある。マニュアルから、評価人材（評価委員会スタッフ等）のマネジメントに係わる事項である、評価の専門家の公募・選定基準・データベース管理・契約等、専門家評価委員（ピア・レビュー）の権限と行動規則、利益相反の宣誓様式、評価パネルのモニタリングのための独立オブザーバーを扱ったものを抜粋した。

また、EU ではプロポーザルを評価する専門家を公募して評価人材を拡充しているが、そのピア・レビューの登録申込み募集の文書を収録した。

B-3.1 科学技術ピアの公募・選定・行動基準・モニタリング等（EU 第 5 次フレームワーク・プログラムの提案公募マネジメントプロセス説明文書）

「プロポーザル評価手続きマニュアル」の構成

イントロダクション	
プロポーザル評価及びプロジェクト選択過程の各段階の概要	annex A 事前プロポーザル・チェックのフィードバックフォーム
事前プロポーザル・チェック	annex B 中小企業固有の手段
事前登録	▶ annex C 専門評価委員の権限と行動規則
プロポーザルの受け付け	▶ annex D 利益相反の宣誓
適格性に関する事務的チェック	▶ annex E 評価に於ける委員会スタッフの役割
▶ 評価の専門家	▶ annex F 評価過程に於ける独立オブザーバーの権限
▶ 評価基準	annex G 評価概要報告書のモデル
プロポーザルの採点	
▶ 独立オブザーバー	以下略
最終審査・プロポーザル・ランク付け・非採択の決定	
契約の作製と締結	
報告	

「▶」は、本資料に記載した評価人材（評価委員会スタッフ等）のマネジメントに係わる事項

▶ 評価の専門家

適格性基準を満たす全てのプロポーザルは、それらの質（評価基準については下記参照）を決定するために評価される。プログラムは独立した外部の専門家を使って、プロポーザルを評価する際に委員会にアドバイスを行う。一般的なルールとしては、最低3名^(注1)の独立した専門家が委員会に提出されたそれぞれ適格であったプロポーザルを審査する^(注2)。

プロポーザルを審査する独立した専門家は、Official Journal of the European Communitiesで公開される公募を通じて探し、当該フレームワークプログラムの期間中オープンにする。公募では、専門家を選ぶための基準を詳しく記述する。一般論として、専門家は第5次フレームワークプログラムに於けるアクティビティの領域で、適切な能力があるものと期待される。彼らはまた、以下の領域またはアクティビティの少なくとも一つについて、公または民間部門で高いレベルのプロ経験を持たなくてはならない。具体的にはつまり、関連する科学・技術分野の研究、RTDプロジェクトの管理・指導・評価、研究開発プロジェクトや技術移転・イノベーションの結果を利用した経験、科学・技術における国際協力、人的資本の開発、である。委員会はまた、候補が、研究の難しさや、産業的あるいは社会経済的な効果を理解する能力、とりわけ、委員会ポリシーと、ポリシーに対するプロポーザルの意味を判断する能力について考慮する。専門家はまた、適切な言語スキルを持たなければならない。専門家として参加するためのセレクションへの申請は、履歴書と、それに基づいてセレクションがなされるであろう適正なキーワードを含む書式による。

専門家パネルのメンバーを選ぶもとなるリストは、以上のような選定基準を用いて委員会スタッフが作製する。専門家のデータベースは、センターが管理して、専門家の適正な入れ替えをきちんとする。

一般に、6年間の任意の期間に於いて、専門家委員は1プログラムについて3年（または3回の公募、のいずれか長い方）以上に渡って務めてはならない。プログラム管理者は、プログラムの一年毎、または一回の公募（一年以上開けて行われる場合）毎に、少なくとも1/3の専門家を入れ替えるようにしなければならない。

公募に応じて申請されるプロポーザルを評価するために、委員会スタッフは、適切な専門家のリスト（必要なら補欠リストを含む）を、適格な専門家のデータベースから作製する。この際、選定されるそれぞれの専門家パネルが、その能力、地理的なバックグラウンド、言語能力に関して、適切な範囲をカバーし、またバランスするように留意する。出来る限り、適切な男女比のバランスをとるようにも留意しなくてはならない^(注3)。更に、理由はなんであれ、これから評価

注1 例えば、Marie Curie フェローシップのプロポーザルは、その性質（比較的小きなプロジェクトで数が多い）からして、最低2名の専門家がプロポーザル評価を行えば良い。

注2 バブリックな調達の手続きを通じて提出された基準や特別助成金についても、この委員会のこのような基準の評価についての通常のルールは適用される。助成金に関する固有の基準は応募者への規定に関する文献10(2)に示されている。随伴基準、特にサイズの小さいものの場合、外部専門委員のアシストなしで評価を実行することを選ぶこともできる。

注3 欧州共同体機会均等ポリシーを追及する。この文脈に於いて、女性はとりわけ、プロポーザルを申請するか、またはその申請に参加するように推奨される。欧州共同体の「女性と科学」の通知では、委員会は、評価パネルの構成に、可能ならば40%女性を含むことを目標として定めることになっている。

することになっている特定のプロポーザルに対して、賛否いずれにせよ偏った意見を持つ可能性のある専門家を招聘しないように留意する必要がある。いずれの評価委員会で採用される専門家のリストも、関連するDirector(s)によって決定される。委員会継続中に必要とされる専門家の交代や追加も同様である。

評価に参加する専門家は、委員会との契約にサインしなくてはならない。これは、専門家委員が調査するプロポーザルに関しての守秘義務と公平性を保証するものである。この契約の付記事項である専門家委員の権限と行動規則については、Annex Cを参照。専門家委員が署名するべき、公平性に関する宣言はAnnex Dに示されている。

大部分のプログラムと公募について、評価を実行するために専門家委員は中心地に招聘されることになる。しかしながら、時間が許す場合、また、商業的なセンシティブティが制限因子でない場合には、プロポーザルは専門家委員に送付（電子的にでも、ハードコピー版でも）されても良く、その場合には各委員は最初の評価を彼らのオフィスで行うことができる。この後者の方法を用いる場合には、そのことは、当該公募の公開されている補遺中で明らかにされる。

この場合、最初の評価の結果は、電子的にか郵送によって委員会に伝えられ、必要ならば、委員会への最終報告を作製するためにパネル内で更に議論される。このような場合、委員会は全ての専門委員を、委員会報告のための議論に招聘することができる。

▶ 評価基準

多くの評価基準は、第5次フレームワークプログラムの全てのプログラムに共通である^(注4)。Annex Iには、フレームワークプログラムについての委員会決定、受託者・研究センター・大学の参加のためのルールに関する委員会決定、間接部門のための選定基準をとりあつかう部門の下で行われる研究結果普及^(注5)のためのルール（「参加規則」）に関する委員会決定、が示されている。それぞれの適格なプロポーザルを、専門家委員がこれらの評価基準に照らして審査する。それぞれのプログラムでは、より詳しい基準を決定し、また、当該プログラムにのみ適用される評価基準を追加できる。評価で使用される基準の個々の説明、及び、基準に適用される重み付けや最低基準はいずれも、この書類の個々のプログラムについてのAnnexに示されており、また、それぞれの公募及び全ての関連する補助資料で参照できる。

参加規則に示されている評価基準に従ってプロポーザルを詳細に審査するために、専門家委員は通常、採点とコメントを作製する。それに加えて、専門家委員は、公募で言及されるスペックに関連した一連の質問に答えることで、特定の評価基準を評価するように要請される。以下の質問は、評価中の適切な時点で検討される。

注⁴ 研究訓練活動(Marie Curie フェローシップ、Human Potential Research Training Networks、Enhancing Access to Research Infrastructures and High Level Scientific Conference)のためには、これら活動に於ける訓練の性格を考慮した固有の基準が適用される。この固有の基準については Annex Nに概略を示しておく。プログラム「イノベーション促進と中小企業の参加奨励」については、これに適応する委員会決定の Annex III-2に、以下のように記されている。パイロットアクティビティについて調整を行う場合、間接的 RTD 行動に一般的に適合するような手続きで始めればよい。但し、公平性と透明性の原則に充分配慮した評価・選定手続きを条件とする。

注⁵ O.J.N L26 of 1.2.99、 pp.46 and 56

- プロポーザルは、ポリシーの問題を含めて当該の公募で公開された作業プログラムのそれぞれの部分を検討しているか？もしプロポーザルが公募案件に部分的にしか適合していない場合、それは、全面的にあるいは部分的に、考慮に値するような十分なメリットを有しているか？
- プロポーザル準備の段階で、関連する倫理的な問題は適切に考慮されているか？即ち、関連がある場合、提案された研究は基本的な倫理原則に抵触しないか？関連がある場合、提案された研究は共同体の政策に合致しているか？即ち、必要な場合、共同体の政策（例えば環境）に従って適切な安全措置や影響評価が考慮されているか？
- プロポーザルは、表現についての要請事項（特に匿名性の要請）を守っているか？

これらの質問に対して否定的な回答であった場合には、専門家委員は、彼らがそのような回答について弁明するようにコメントしなくてはならない。専門家委員のコメントに基づいて、委員会は、上記のような要請項目の少なくとも一つを満たさないことが明らかとなった任意のプロポーザルの評価を停止する権利がある。明確なケース（例えば当該公募で行うと言っていないような研究課題についてのプロポーザル）としては、プロポーザルが課題の範囲から全く外れていたり、適格性チェックが行われた時点で明記されているポリシー要請に反していたりする。

公募課題に部分的にしか合致していないプロポーザルや、複数のプログラムの領域にまたがったプロポーザルはケースバイケースで審査される。ケースとしては以下のようなものがあり得るだろう。そのプロポーザルを全体として評価する場合もあるし、プロポーザルの内容のうちで公募課題に合致している部分だけを評価する場合もある。あるいは、（ちょうど良い別のプログラムが公募している場合）別のプログラムへ回す場合もあるし、公募課題に合致している部分が当該プロポーザルのうちではかなりマイナーで、その部分を他から切り離すとあまりにつまらなくなってしまうような場合には、当該プロポーザルの評価を諦めて辞める場合もある。公募課題の要請と合致している全ての適格なプロポーザルは、それらの質と妥当性を外部専門家のアシストを得て委員会が審査する。このマニュアルのプログラム固有のannexで示されない限り、専門家委員はプロポーザルを審査し、採点する。評価基準は以下に示す通り（これらはフレームワークプログラムの決定と「参加規則」の決定に従って作製されたリストであり、大きく5項目に分類されている）。加えて、専門家委員はそれぞれの評価項目について総合点をつける（但し、プロポーザルがいずれかの最低基準値を下回らない限り。以下参照）。専門家委員は、申請者にフィードバックを提供できるような書式で、採点のそれぞれについてコメントをつける。これらのコメントは採点結果と矛盾しないようなものでなくてはならない。

全てのプログラムで用いられる評価基準の項目は以下の通り。

科学・技術的な質とイノベーション

- プログラムまたはキーアクションの目的を達成する、との観点による、提案課題の質、及び、科学・技術的重要課題の研究への貢献度。
- オリジナリティー、イノベーションの度合い、現状の技術水準を越えた進歩、プロジェクトに関連するリスクへの配慮。
- 選択したアプローチの適切さ。科学・技術的目標を達成するための方法論と作業プラン。

(欧州) 共同体付加価値とEU政策への貢献

- 課題の全ヨーロッパ的広がり。ヨーロッパレベルの問題解決に貢献するプロジェクトの広がり、及び、ヨーロッパレベルで行われた作業の期待されるインパクトが国別プロジェクトの「合計」よりも大きいような広がり。
- コンソーシアムのヨーロッパ的付加価値。全ヨーロッパ規模での、異なる組織に於ける利用可能な資源や相補的な専門知識の組み合わせ、及び、人的・財政的な意味での臨界量を確立する必要性。
- 一つ以上のEU政策（小・中規模企業（SMEs）向けの政策などのような、「水平」政策）の実行または発展に対するプロジェクトの貢献。または、標準化や規制に関連する問題の検討。

(欧州) 共同体社会的な目的に対する貢献

- 生活・健康・安全の質（労働条件を含む）の改善に対するプロジェクトの貢献。
- ヨーロッパに於ける雇用見通し、スキルの発展と使用の改善に対するプロジェクトの貢献。
- 環境の保全または増進、及び、自然資源の保存または最少利用へのプロジェクトの貢献。

経済発展と科学・技術的見通し

- 成長、特に有用性と応用の広がり、開発計画の質（提案プロジェクトに基づくRTDの結果に対する開発活動を行うパートナーの信頼性を含む）、プロジェクトのより広い経済的インパクトへの、考えられる貢献。
- 提案プロジェクトの戦略的インパクト。及び、それが、RTD結果のユーザーやパートナーにとってのアプリケーション市場の開発や競争状況の改善に寄与するポテンシャル。
- ヨーロッパの技術的進歩への貢献。とりわけ、期待される結果の普及戦略、ターゲットグループの選択など。

資源・パートナーシップ・マネジメント

- 提案されたプロジェクトの方法論、及び管理の質。とりわけ、適正さ、明快さ、整合性、提案されたタスクの効率と完成度、進行計画の調整（マイルストーンがあること）、マネジメントの構成。加えて、インパクトとパフォーマンスの評価軸として示されたものの質を含む、プロジェクト進行をモニターするツール。及び、プロジェクトコンソーシアムの内部でコミュニケーションを良好な状態に確保すること。
- パートナーシップの質。適正ならば、ユーザーや当該分野内での他のメンバーの参加。とりわけ、コンソーシアム内部での科学・技術的能力と専門知識及び役割と機能。パートナーとの相補性。
- 資源の適正さ。それぞれのパートナーあるいはタスクへ振り分けられるマンパワー。振り分けられるマンパワー、耐久財、消耗品、旅費、及びその他の使用されるあらゆる資源の質またはレベルまたはタイプ。加えて、予算に組み込まれない資源（例えば、研究設備、キーパーソンの専門知識）。この評価項目については、採点よりも寧ろコメントが与えられるべきであろう。

上記で概説した評価基準に加えて、もしプロポーザルがフレームワークプログラムに参加しない第三国の参加者を含む場合には、専門家委員は、それら参加者の参加が共同体の利害と一致するかどうか、また、プログラムの目的に照らして当該プログラムの全体または部分の遂行に実質

的な付加価値を与えるものかどうかを審査する。もしそのような参加者が共同体のファンドを要請するならば、専門家委員はそのような資金提供が当該問題の目的の達成に不可欠かどうかを審査する。

プロポーザル審査にあたっては、専門家委員はこのマニュアル中で示された評価基準だけを適用する。但し、プログラム決定に含まれる任意のプログラム固有の評価基準が補遺として追加され得る。これらの特定プログラムに適用されるような評価基準は、プログラム固有のannexと作業プログラムにかなり詳細に記述されているであろう。専門家委員が、このマニュアル及びプログラム固有のannexで提示されている評価基準から逸脱した評価基準を適用することは許されていない。

必要に応じて（例えば、相当の応募超過が予想される場合や新たな複雑な研究テーマが導入される場合）、公募は2段階（2回申請する）に分けて行われても良い。第一回の申請（これは簡単な書式による）の後、評価は申請研究の質とイノベーションの度合いにしばって行われる。この第一段階の評価によって、委員会は、要求レベルに達していない「概要」申請の非採択を決定する。この段階で非採択されなかったプロポーザルのコーディネーターは、評価の第二段階のために、一定の期間のうちにフルプロポーザルを提出するように連絡を受ける。この場合、このマニュアルで示した評価のルールの本質的な部分は、評価プロセスの全体にわたって適用される。それぞれのステップに於ける評価基準の内訳や重み付けについては、プログラム固有のannexで明示され、また、公募説明とその補遺の中で言及される。

評価委員会は全体的に委員会スタッフに監督される。彼らの役割については Annex E に示した。

▶ 独立オブザーバー

年次プログラムモニタリングに併せて、また、それぞれのプログラムについて年に一回以上、評価委員会をモニターするために独立オブザーバーが招聘される。彼らの任務は、このマニュアルで、あるいは個々のプログラムによって公開されている補遺情報に示されている手続きが守られているかどうかを確認すること、及び、改善できるプロセスのやり方をプログラム管理者に報告することである。2段階申請方式のプロポーザルの場合、オブザーバーは、いずれの段階も含めて全プロセスに関する手続きの適用について報告する。オブザーバーは、モニタリング公募にアプライした者、評価の専門家委員、RTDプログラムに関する専門家アドバイザーの中から選ばれる。独立オブザーバーの権限については Annex F に示す。

【 Annex C 専門家評価委員の権限と行動規則 】

1. 専門家委員のタスクは、このマニュアルとそれぞれのプログラム固有の評価基準ドキュメントに記述された手続きに従って、それぞれのプロポーザルの、守秘義務を守り、公正で公平な評価活動に参加することである。専門家評価委員は、その最善をつくしてこれを達成する。また、委員会スタッフによるあらゆる指示に従い、恒常的で高くオリティの作業を行う。
2. 専門家委員は、委員会との契約の下、独立した個人として作業にかかわる。専門家委員は個人的能力に従って作業すると想定され、その作業遂行中には、たとえ報酬支払い契約が当該専門家委員の特定の組織における雇用関係を示しているとしても、他のいかなる組織も代表することはない。
3. 専門家委員は、作業を始める前に守秘義務宣誓に署名する。そうすることによって、専門家委員はその作業にあたって自分自身を厳格な守秘義務と公平性に専心させるのである。当該宣誓に署名しない招聘された専門家委員は、評価委員として作業に参加することは許されない。専門家委員が、直接にせよ間接にせよあるプロポーザルに関係がある場合、あるいはいかなる他に帰属する利害関係がある場合、即ちなんらかの方法でプロポーザルに繋がりがあある場合、または何らかの他への忠誠心が、プロポーザルに関する委員の公平性を損ねる場合、あるいは損ねるおそれのある場合には、当該委員は、そのことに気付いた後できるかぎり速やかに、責任を有する委員会スタッフにその事実を申告しなくてはならない。その繋がりがそのようにして専門家委員の公平性を損ねるおそれがあるほどに強い場合には、委員会スタッフは、当該専門家委員の当該プロポーザル、及び必要に応じて競合プロポーザルの評価への不参加を確保する。

専門家委員は、以下のような場合にはあるプロポーザルと直接的なリンクがあると想定される。

- 専門家委員が、現在、または最近、申請している組織から雇用者を得た場合。
- 専門家委員が、当該プロポーザルの作製に参加している場合。
- 専門家委員が、プロポーザル申請チームの申請者またはそのメンバーと関係がある場合。
- 専門家委員が、当該の結果の出版や宣伝にそれと知りながら参加しているかも知れない場合。

専門家委員が、当該プロポーザルがカバーする分野で、申請組織のうちの一つと契約上の関係を持つような組織に雇用されている場合、あるいは、専門家委員が、競合プロポーザルを申請している組織と直接の関係がある、またはその組織で働いている場合に、当該委員はそのプロポーザルと間接的なリンクがあるものと見なされる。

4. 専門家委員は、責任のある委員会スタッフが認識し調整したか、当該スタッフが議長を務める会議において正式な議論を行うその時を除いては、当該プロポーザルの評価に直接は参加していない他の専門家委員や委員会スタッフを含めて、いかなる他人とも、どのプロポーザルに関しても議論してはならない。
5. 専門家委員は、申請者と連絡を取ってはならない。また、評価期間を通じて全てのプロポーザルは改訂されてはならない。すべてのプロポーザルについての委員会への専門家委員の勧告について、専門家委員が、申請者に対しても、それ以外の誰に対しても連絡するようなことはあってはならない。
6. 専門家委員は、評価に参加している他の専門家委員の名前を漏らしてはならない。委員会事務局が、誰がどのプロポーザルを評価したかを示すことなく、専門家委員の名簿を決まった期間の後に公表する。
7. 専門家委員に対して、プロポーザルが電子的に送付、または郵送されることが決まった時、普段作業している場所であれ他の適当な建物であれ、当該専門家委員は、送付された全ての書類あるいは電子ファイルに関する守秘義務を維持する責任を有し、また、評価終了直後に、守秘対象の書類やファイルを消

去あるいは破壊する責任を有する。そのような場合、専門家委員は、当該プロポーザルの審査を完了できるように、更なるアドバイスや情報を求めてかまわない。他人との議論や接触が、守秘義務と公平性に関する総合的な規定を遵守している限りにおいて。

8. 委員会が管理しているオフィスや建物で評価が行われる場合には、専門家委員は、プロポーザルのどの部分も、メモのコピーも、それが紙媒体であれ電子媒体であれ、プロポーザルの評価に関連するものは何であれその評価を行っている建物の外に持ち出すことは許されない。プロポーザルに関する全ての情報は、その評価を行っている建物から専門家委員が出かけている間は、セキュリティを保って保管される。
9. 評価を行っている構内では、専門家委員は、評価が始まる時に渡されるバッジを見えるように着用していなくてはならない。バッジがない時には、関連する委員会スタッフの特別な許可がなければ、構内に入ることは許されない。専門家委員は、契約の最終日にそこを立ち去る際に、バッジを評価組織に返却する。
10. 専門家委員は評価プロセスの守秘義務を確保するために委員会事務局が定めるあらゆる規則(例えば、評価委員会の外で人とコミュニケーションすることに関する規則)に厳格に従うことが要求される。もしもこの規則遵守が破られた時には、それ以上の評価プロセスからすぐさまに排除される。

【 Annex D 利益相反の宣誓 】

(何れか該当する項目があればチェックして下さい)

下記に署名したわたくしは、専門家評価委員の権限と行動規則を読み、理解したことを確認します。

わたくしは、 _____ の公募の下に評価のため、自分自身がプロポーザルを提出していません。また、提出されたいずれのプロポーザルについても、わたくしの知る限り、直接または間接に関係していることはありません。以上、申告致します。

わたくしは、以下のプロポーザルの評価にわたくしが参加する場合、直接または間接に利害関係が生じる可能性があることを申告致します。

分野の頭字語

タイトル

私は、私が評価を依頼された、あるいは私が出席するすべての評価委員会会議において議題となるすべてのプロポーザルに関して、直接にせよ間接にせよ、利害関係があることが判明した場合には、ただちに委員会スタッフに報告することを確約いたします。

サイン _____

名前 _____

日付 _____

【 Annex E 評価に於ける委員会スタッフの役割 】

1. 委員会スタッフはプログラム固有の補遺に記述された評価基準に従い、それぞれのプロポーザルの評価の守秘義務・公正性・公平性を組織する。また、このタスクのために明記された妥当な手続き・ルール及び規制の遵守して申請者をガイドする。
2. 委員会スタッフは、評価のためにプロポーザルを外部専門家委員に割り当てる。これを行うために、プロポーザル（または競合しているプロポーザル）が直接・間接に当該プロポーザルと関係を持つような専門家委員に割り当てないように留意しなくてはならない（このような関係の定義についてはAnnex Cを参照）。
3. プロポーザル評価のための専門委員選定にあたっては、委員会スタッフは以下のことに更に留意する必要がある：
 - 必要な能力の適切な範囲
 - アカデミックソサエティに属する専門家・産業界の専門家・ユーザーの間の適正なバランス
 - 無理のない男女比^(注6)
 - 専門家委員の出身地に関する無理のない分布
 - 評価間の専門家委員の定期的な入れ替え。
4. 委員会スタッフは、必要ならば、プロセスの適正な実施を確実にするように行動する。これは以下の行為を含む。即ち、従うべき手続きについての専門家委員への説明、専門家委員への規則の徹底、委員会が正式に責任を負うべき全ての不規則要件の報告、遵守すべき契約事項や守秘義務に違反していると考えられる人物のプロセスからの排除。任意の人物の排除が行われた場合には、組織単位の責任者またはプログラム責任者に適正に報告する。
5. 委員会事務局への勧告を確立するために専門家委員会会議をコーディネートする場合、委員会スタッフは議長役を果たし、特定のプロポーザルや関連組織へのいかなる賛否の先入観を排除して、専門家委員間の合意形成を求める。評価パネルの会議に出席する委員会スタッフは、プロポーザルの適正な評価を行うために必要となるあらゆる追加説明・情報を供給する。
6. 委員会スタッフは、専門家委員会の作業のパフォーマンスを監督する責任を有する。そこでは上記で言及されてきたような要件が考慮されているかどうかチェックしなくてはならない。
7. 委員会スタッフは、「監査記録」（即ち、それぞれのプロポーザルが持つべき全てのファイル、例えば、専門家委員の採点表やコメント）を維持する責任がある。個々の専門家委員の採点表から採点記録を複製し、任意のプログラム固有の評価書類に明記されたルールに従って、合意形成に至るのに必要な議論のあらゆる基準を明確にしなくてはならない。
8. 委員会スタッフは、それがプログラム責任者あるいは組織単位の責任者によって適正に明示的に認定されている（ケースバイケースで決まる）場合でない限り、評価や選定プロセスに於ける論点について、申請者やプロセスに直接関係のないどんな人とも議論してはならない。このようなことは、例外的な場合、プロセスの守秘義務遵守の必要性について充分考慮された場合にしか発生しない。
9. 委員会スタッフは、専門家委員へのプロポーザルの割り当てについて、厳格に守秘義務を遵守する。評価に参加した全ての専門家委員のリストは、具体的な割り当ての情報は示さない形で、通常の間隔で公

注⁶ 注6参照

開される。

10. 委員会スタッフは、プロポーザル及び評価に関連する全ての書類に関する適正な守秘義務措置を確実にするための全ての必要な手段を取る。とりわけ、
 - プロポーザルと関連書類は、申請者が明示的に同意しない限り、作業の適正な遂行上必要とする委員会役員、専門家委員、申請者本人、以外のいかなる人物にも見せてはならない。
 - 専門家委員による、委員会への評価報告と勧告は、作業の適正な遂行上必要とする委員会役員、及び、当該評価プロセスに割り当てられたオブザーバーまたは監査委員だけに閲覧は限定されている。
11. 委員会役員は、プロポーザル及び評価ドキュメントの複写を制限する。また、もはやそれが必要でなくなった時には、評価期間に用いられたコピーや書類及びメモの類を確実に破壊しなくてはならない。

【 Annex F 評価過程に於ける独立オブザーバーの権限 】

オブザーバーの役割は、評価委員会の遂行について、改善できるであろう手続きの方法について、会議で用いられている評価基準の適正さについて、及び、専門家委員によるこれら評価基準の適用方法について、独立な勧告を委員会に対して行うことである。

オブザーバーは、一般的な責任者または当該案件責任者が、場合によって、年次のモニタリング活動をそれぞれのプログラムについて行っている高レベルの専門家委員間で、あるいはまた、モニタリング・評価専門委員・専門家アドバイザーの公募に応募したところのある者の中から選任する。このような選任は、様々な活動の間の連続性をモニターすること、及び、とりわけ、監視パネルが直接の経験から評価プロセスについて確実にコメントできるようにすることを目的としている。オブザーバーは、評価プロセスを、その結果ではなくて、それがきちんと機能しているかどうかと云う観点から審査することを求められる。特に、オブザーバーは審査中のプロポーザルについての見解や、プロポーザルに対する専門家委員の意見についての見解を表明しない。

オブザーバーは、専門家委員が委員会スタッフに状況説明を受ける、最初の評価委員会の席に同席するように招聘され、委員会が規定した評価基準に関する視点での合意形成がなされるプロセスや議論を監視するために、評価パネルを訪問するように招聘される。

オブザーバーは、当該案件責任者に彼らの発見を報告し、適正な一般的な責任者に報告書のコピーを送る。彼らの報告のサマリーは、公に利用可能な形で公開される。レポート作成と同様にして、オブザーバーはまた、評価委員会中の委員会スタッフとの非公式な議論に参加することを、また、プログラム責任者・プログラム管理者・当該評価に責任を有する役員がすぐさま実行に移せるような可能な改善を見つけるべく注意深い監視を行うことを推奨される。

オブザーバーは、専門家委員と同じ守秘義務 (Annex C) を遵守することが要請され、守秘条項同意書に署名する。プロポーザルの詳細や、プロポーザル評価に割り当てられている専門家委員を漏らしたり、評価パネルにおける議論を漏らしたりすることは許されていない。

B-3.2 科学技術ピアの公募文書

「特定研究プログラムに関して受け取るプロポーザル評価のための専門家リスト登録申込み募集」

1. ヨーロッパ共同体研究およびテクノロジー開発、デモンストレーション活動第 5 フレームワークプログラムに対するプロポーザル¹、ならびに、ヨーロッパ原子力共同体研究およびトレーニング活動第 5 フレームワークに対するプロポーザル²、第 5 フレームワークプログラム実施に向けての規則に関する評議会決定に対するプロポーザル³の目標を達成するため、上記のフレームワークプログラムを実施する特定研究プログラムに関連する応募として受け取ったプロポーザルを評価する専門家リストを作成することを目的として、委員会は公募を行う。

第 5 フレームワークプログラムの主要な目標のひとつは、主に経済的・社会的問題の解決に関して、共同体の研究がより戦略的・指導的役割を中心となって果たすことである。複雑な学際的問題を効果的に解くためには、資金提供された研究分野およびプロジェクトに向けて統合的アプローチが必要である。

2. 委員会が専門家パネルへの登録者として選択した応募者は、ポイント 1 で言及した文書に述べられたさまざまな科学上・テクノロジー上・社会経済上の目標——国際協力も含む——に沿ってプロポーザルを評価することを支援する。評価作業は、問題の特定プログラムの目標を最大限達成するよう研究を適応させる仕方に関する勧告を作成することも含まれる。

第 4 フレームワークプログラム実施のために作成された専門家リストに既にあり、第 5 フレームワークプログラムのもとプロポーザル・アセスメントへの参加を希望する応募者は、あらたに応募書類を提出する必要がある。

3. 加盟国、もしくはフレームワークプログラムに参加する国、委員会が科学・テクノロジー分野の協力合意の締結国としてきた国の国民は、応募資格がある。他国の国民からの応募は受理されるが、各特定プログラムの必要に応じてのみ使用される。

ポイント 1 で言及した文書および応募様式で示されているように、応募者は第 5 フレームワークプログラムへのプロポーザルで述べられる活動分野において適切な能力を有する証拠を提出しなければならない。同様に、次の分野もしくは活動のひとつ以上において、私的セクターもしくは公的セクターでの高いレベルの専門経験を有さねばならない。

- a) 妥当な科学・テクノロジー分野における研究
- b) 研究・テクノロジー開発プロジェクトの管理もしくはマネジメント、評価
- c) 研究・テクノロジー開発プロジェクトの成果の活用、テクノロジーおよびイノベーションの移転
- d) 科学・テクノロジーにおける国際協力
- e) 人材開発

同様に、委員会は、とくに共同体の政策という点から研究のチャレンジおよび、産業上および／もしくは社会経済上の影響を評価する応募者の能力も考慮する。

4. 応募書類は以下に述べる手続ののって提出されなければならない。
応募様式は次のウェブサイト・アドレスから入手可能である。

<http://www.cordis.lu/expert-candidature>

また、次に請求してもよい。

European Commission

DG XII – SDME 5/20

200 rue de la Loi

B – 1049 Brussels

応募様式は、ヨーロッパ連合の公式言語のひとつで記入されなければならない。

応募書類は次のいずれかの方法によって送付しなければならない。

- ・受信通知付きの書留によって次の住所へ送付する。

European Commission

DG XII – SDME 5/20

200 rue de la Loi

B – 1049 Brussels

- ・受信通知に対する返信として、宅配便もしくは本人によって、次のアドレスへ送付する。

European Commission

DG XII – SDME 00/33

8 square de Meeus

B – 1050 Brussels

封筒には次のような表書きを添えなければならない。” Call for applications for expert evaluators No...” (専門家評価公募 No...)

応募書類は e-mail によって提出してもよい。 e-mail 応募の送信開始日および e-mail アドレスはヨーロッパ共同体オフィシャルジャーナル (Official Journal of the European Communities) に掲載され、また、ほかの適切な手段によって周知される。

5. 受理された各応募書類は、応募様式中のキーワードを基礎にすべての特定プログラムを対象とする潜在的専門家リストに載せられる。このリストは、2002年9月30日までの評価パネルを編成するために用いられる可能性がある。

委員会は、パネルがバランスの取れた仕方設定され、専門家の適切なローテーションが行われることを保証する。委員会は、応募者の地理的出自や専門的背景 (産業および行政、研究およびイノベーション、社会経済的ユーザーおよびセクター) を考慮に入れる。同様に、委員会は女性と男性のバランスの取れた参加を求めている。

1999年2月1日以前に受理された応募書類のみが、各特定プログラムによる第一次プロポーザル募集に対して応募されたプロポーザルを評価する専門家パネルへの登録を対象とすると見なされる。1999年2月1日以降に受理された応募書類は、それ以外のプロポーザル募集のアセスメントを対象とするとみなされる。

パネルとして選ばれた専門家のリストは、厳正な手続のもと公刊される。

6. プロポーザルの評価の独立性を保証するため、契約の終わりに、選ばれた専門家は自分自身が選ばれた評価パネルの作業と自分が就いている地位との間に利害衝突がないことを確言する宣誓に署名しなければならない。評価プロセスを通じて、専門家は義務に対する適切な献身を目に見える形で示し、プロセスにおいて自分自身の注意にさらされた情報と文書

の機密を守らなければならない。

7. 専門家契約は、選ばれた専門家、もしくは専門化が法人によって雇用されている場合には、その法人の正当な代表者のいずれかが署名しなければならない。選ばれた専門家は、契約の署名の時点以後、妥当な率（scale）にしたがって報酬が支払われる。旅費および実費は、委員会内で実施されている規定をもとに補償される。

註：

- 1) 情報は以下で入手可能。 <http://www.cordis.lu/fifth>
- 2) OJ No C 106, 6.4.1998, p.45.
- 3) OJ No C 262, 19.8.1998, p.50.

（平成 12 年度「外部評価機関の在り方に関する調査報告書」（財）政策科学研究所 より抄録）

参考資料 2.

欧米の評価研究者・エキスパートとネットワーク

本調査では、評価人材の研修コースの発展的なものとして、海外の実践事例や研究事例に学ぶエキスパート・コース等の検討を行っている。海外の先行的な事例や研究成果は、我が国の評価システムの改善にとって、環境的文脈的な相違、とくに文化的な相違に留意する必要はあるが、示唆される内容が豊富な対象である。また、本文中にも述べたが、近年拡充が進んだ主要な研究者ネットワークや国際会議に参画している研究者や実務家は、研修コースでの海外講師陣の対象母集団である。本資料(2.1)には、主要な国際的な研究者ネットワークとその発展過程、国際会議等の成果物や出版物、属している専門的研究者のリストやホームページ等を採録した。なお、本調査では、とくに米国の行政機関内の実務家に関して、研修講師候補を調査したが、その連絡先等を記したリストを(2.2)に付した。

2.1 評価研究者・エキスパートの国際的ネットワーク

2.1.1 The World Research Evaluation Network

<http://www.reseval.net>

<http://www.prism.gatech.edu/~sc149/reseval/index.html>

◆ 目的と活動の概要

世界の評価研究者と評価機関 (offices) が経験、方法、データを共有するための組織である。この web サイトには、新しく公表された論文の紹介、今後開催される会議の広報や過去の会議の記録の掲載を行っている。かなり大量の近年の重要文献資料が蓄積されている。

活動の目標は、研究評価 (research assessment) の手法に関する最新状況をユーザーのコンテキストを含みつつレビューすること、そしてその分野の研究のベースとなるアジェンダや評価方法の開発について概観することである。

◆ 活動の内容

The Res. Eval. discussion list という名前のメーリングリスト(購読制)を運営しており、世界中の研究評価者にリンクできるようになっている。サイトのアドバイザー・ボードとして、Phillipe Laredo, Philip Shapira, and Gretchen Jordan が関与し、サイト自体のスポンサーは NSF の Program of Research on Science and Technology である。Susan Cozzens (susan.cozzens@pubpolicy.gatech.edu.) がサイトの更新にあたっている。

またこのネットワークのウェブサイトには、関連する世界中の研究者、雑誌、研究機関、コンサルタント会社、政府機関へのリンクが張られている。

◆ ウェブサイトに掲載（リンク）されている資料

- ① 研究者および研究グループ／助成機関等の評価グループ
- ② 国際会議情報
- ④ 評価方法に関する重要文献（ダウンロード可能）
- ⑤ 当ネットワークが関与した評価活動ならびに評価研究活動（ダウンロード可能）
- ⑥ 評価に関連する情報源（S&T データ、連邦政府のプログラム評価関連情報源）
- ⑦ 専門誌“Research Evaluation”について

◆ 研究者および研究グループ

▶ Asia / Pacific

- Ron Johnston (University of Sydney
ACIIC, Australian Centre for Innovation and International
Compettiveness, Australia)
- Tim Turpin (University of Wollongong, Centre for Research Policy, Australia)

▶ Canada

- Benoit Godin (Universite du Quebec a Montreal
CIRST, Centre Interuniversitaire de Recherche sur la Science et
la Technologie)
- Yves Gingras (Universite du Quebec a Montreal, CIRST)
- Benlit Godin (OST, Observatoire des sciences et des technologies)
- Adam Holbrook (Simon Fraser University
CPROST, Centre for Policy Research on Science and
Technology)

▶ Europe

- David F.J.Campbell (IHS, Institute for Advanced Studies, Austria)
- Stefan Kuhlmann
(ISI, Fraunhofer Institute for Systems and innovation Research, Germany)
- Research Policy Institute (Lund University, Sweden)
- Luke Georghiou (University of Manchester,
PREST, Public Research in Engineering, Science and Technology)
- Ben Martin (University of Sussex at Brighton,
SPRU, Science and Technology Policy Research)
- 個人名記載なし (Evaluation Associates Ltd.)
- Erik Arnold (Technopolis, Innovation Policy Research Associates)
- Patries Boekholt (Technopolis, Innovation Policy Research Associates)
- Pirjo Niskanen (VTT, Group for Technology Studies, Finland)
- Terttu Luukkonen (VTT, Group for Technology Studies, Finland) (注) 現籍は異なる。
- Juan Campanario (Universidad de Alcala de Gebares, Spain)

▶ Latin America

- Judith Sutz (Republic University of Uruguay
CSIC, Comision Sectorial de Investigacion Cientifica)

- Sandra de Negraes Brisolla (State University of Campinas, UNICAMP Brazil
Geosciences Institute
DPCT, Department of Scientific and Technological Policy)
- Lea Velho (State University of Campinas, UNICAMP, Brazil)
- ▶ United States
 - Ronald Kostoff (Office of Naval Research)
 - Barry Bozeman (Georgia Institute of Technology,
RVM, The Research Value Mapping Program)
 - Juan Rogers (Georgia Institute of Technology, RVM)
 - Susan Cozzens (Georgia Institute of Technology
TPAC, Technology Policy & Assessment Center)
 - Gordon Kingsley (Georgia Institute of Technology, TPAC)
 - Alan Porter (Georgia Institute of Technology, TPAC)
 - David Roessner (Georgia Institute of Technology, TPAC)
 - Philip Shapira (Georgia Institute of Technology, TPAC)
 - Irwin Feller (Penn State University, Institute for Policy Research and Evaluation)
 - Diana Hicks (CHI Research Inc.)
 - Francis Narin (CHI Research Inc.)
 - Michael Crow (CSPO, Center for Science, Policy & Outcomes)

◇ 助成機関等の評価グループ

- ▶ Canada
 - Conseil de la Science et de la Technologie, Gouvernement du Quebec
http://www.cst.gouv.qc.ca/cst_English.html
- ▶ Europe
 - DTI, Department of Trade and Industry, United Kingdom
TESE, Technology, Economics, Statistics & Evaluation Directorate
<http://www.dti.gov.uk/tese> (John Barber)
- ▶ United States
 - NIST, National Institute of Standards and Technology
<http://www.nist.gov>
ATP, Advanced Technology Program/Economic Assessment Office
http://www.atp.nist.gov/eao/eao_main.htm
 - National Science Foundation
<http://www.nsf.gov>
Office of Integrative Activities
<http://www.nsf.gov/od/oia/start.htm> (Nathaniel Pitts)
REC, Research, Evaluation & Communication
<http://www.ehr.nsf.gov/EHR/rec> (Eric Hamilton)
 - U.S. General Accounting Office
<http://www.gao.gov> Robin (M.Nazzaro)

◆ 国際会議情報

- “Seventh International S&T Indicators Conference”, Karlsruhe, Germany (September 25-28 2002 予定)
- “Evidence for Better Decision-Making”, Canadian Evaluation Society, Annual Conference, Halifax, Nova Scotia (May 5-8, 2002 予定)
- “International Conference on Socio-economic Evaluation of Public R&D”, Tokyo, Japan (April 18-19, 2002 予定)
- “Evaluation and Transition: devolution, diversity and evidence based change”, United Kingdom Evaluation Society (UKES) Annual Conference Belfast (December 6-7, 2001)
- “International best practices in evaluation of research in public institutes and universities”, Brussels, Belgium (November 12-13, 2001)
- “Mainstreaming Evaluation, American Association Evaluation (AEA) Annual Conference”, St.Louis, Missouri (November 6-10, 2001)
- “The Fifth Iberoamerican, Interamerican Workshop of Science and Technology Indicators”, Montevideo, Uruguay (October 15-18, 2001)
- “China's Economy: Confronting Restructuring, Stability, and International Competitiveness”, University of Wollongong, Australia International Business Research Institute Department of Economics (July 14-15, 2001)
- “Strengthening the Capacity of Science and Technology Evaluation in APEC Region Science and Technology Evaluation in the 21st century”, (May 30 - June 3, 2001)
- “Powering Up for Evaluation: The Odyssey Continues”, (May 21-23, 2001)
- “Research Assessment: What's Next?” (May 17-20, 2001)
- “EC conference on the contribution of socio-economic research to the benchmarking of RTD policies in Europe”, (March 15-16, 2001)
- “Learning from S&T Policy Evaluation”. (September 11-14, 2000)
- “Evaluation of Science and Technology in the New Europe”,. (June 7-8 1999)
- “Civilian Technology Policy in the European Union and the United States: Recent Experiences and New Directions”, (April 9-10, 1999)

◆ 評価方法に関する重要文献

- Erik Arnold, Katalin Balázs, “Methods in the Evaluation of Publicly Funded Basic Research”, Technopolis, March 1998.
- Philippe Laredo, “The development of a reproducible method for the characterisation of a large set of research collectives”, Armines/CSI, March 1999.
- Ken Guy, Katalin Balazs, James Stroyan, Erik Arnold, “Strategic Options for the Evaluation of the R&D Programmes of the European Union”, Technopolis, November 1998.
- “Policy Evaluation in Innovation and Technology. Towards Best Practices OECD”, Organization for Economic Co-Operation and Development.

◆ 専門誌『Research Evaluation』について (<http://www.scipol.demon.co.uk/re.htm>)

ピアレビューによる国際誌。個々の研究プロジェクトの評価から研究活動の国際比較まで

を幅広く扱う。公立機関も民間機関も対象とし、自然科学も社会科学も扱う。4月、8月、12月の年3回の発行（発行は Beech Tree Publishing）。

- 編集主幹 Professor Anthony van Raan (University of Leiden, The Netherlands)
Susan Cozzens (Georgia Institute of Technology, USA)
- 書評編集主幹: Michael Keenan (University of Manchester, UK)
- アドバイザリーボード: Andy Boddington (Evaluation Associates, UK),
Bruce Campbell (Servier Laboratories, UK),
Luke Georghiou (PREST, Manchester University),
John Holmfield (formerly US House of Representatives),
Ron Johnston (ACIIC, Sydney, Australia),
Wilhelm Krull (Max-Planck Gesellschaft, Germany),
Philippe Laredo (Ecole des Mines/CSI, France),
Terttu Luukkonen (Tech Research Centre of Finland),
C Marciano da Silva (New University of Lisbon),
Ben Martin (SPRU, Sussex University),
Frieder Meyer-Krahmer (Fraunhofer Institute, Germany),
Francis Narin (CHI Research, USA),
Erkki Ormala (S&T, Helsinki, Finland),
Giorgio Sirilli (National Research Council, Italy).

2.1.2 Members of the "6 Countries Programme"

– The International Innovation Network"

<http://www.6cp.net/>

◆ 目的と活動の概要

1975年に設立された、イノベーションの研究と政策形成にかかわる専門家、政策形成担当者、政策実施者などからなる国際的なネットワーク。設立にかかわったのはオランダ、ドイツ、フランス、英国の4カ国だが、現在はメンバーが、オーストリア、ベルギー、カナダ、フィンランド、ハンガリー、アイルランド、スウェーデンと合計11カ国にまたがる。

1年に2回開催される運営委員会 (Steering Committee) には各国から少なくとも1名が入り、ワークショップを企画する。秋と春の2回のワークショップが最も重要なイベントとなる。インフォーマルで開かれた雰囲気の中で、各国の最新情勢に関する情報と意見を交換し、この領域の専門家たちの研究を刺激することをねらいにしている。成果は6CPのウェブサイト、ニューズレター、そして雑誌論文や出版物の形で公表される。

◆ the 6 Countries Programme のメンバー (The International Innovation Network)

- Austria: Mr. Georg Panholzer, Mr. Wolfgang Polt
- Belgium: Mr. Paul Zeeuwts
- Canada: Mr. Geoffrey Nimmo, Dr. Micheal Kelly
- Finland: Mr. Torsti Loikkanen
- France: Mr. Remi Barre
- Germany: Prof. dr. Stefan Kuhlmann

- Hungary: Dr. Andras Siegler
- Ireland: Mr. Dermot O'Doherty, Mr. Gerry Sweeney
- The Netherlands: Mr. Arie van der Zwan, Mr. Jaap van der Vlies, Prof.dr.Ruud Smits
- Sweden: Mr. Lennart Elg
- United Kingdom: Mr. Tony Pedrotti

◆ 最新回・次回のワークショップ

- Innovation policy and sustainable development
Can public innovation incentives make a difference?
Flanders, Six Countries Programme Conference,
Brussels, 28 February- 1 March 2002
- New Governance for Innovation
14-15 November 2002, Karlsruhe, Germany:
- Governance of Innovation: Policies and Strategies within Organisations, Countries
and Regions
Spring 2003, Canada

◇ 過去のワークショップ

49. Merging Innovation Cultures - Leveraging Creativity for Competitive Advantage
(4-5 October 2001, Dublin)
48. Innovation Policies for a New Era (15-16 January 2001, Stockholm)
47. Services in Innovation - The Role of Knowledge Intensive Business Services in the
Innovation Process of Firms (11-12 May 2000: Utrecht)
46. Services Innovation and the Knowledge Economy (22-23 April 1999: Manchester)
45. Knowledge Transfer in the Information Society (19-20 November, 1998: Budapest)
44. R&T Co-operation with Central and eastern European Countries (25-16 May,
1998: Vienna)
43. Local Innovation - New Social, Economic and Technological Dimensions
(November 1997: Dublin)
42. The past and Future of EU Research Funding - Framework programmes under
consideration (16-17 June 1997: Helsinki)
41. Innovation and Sustainable Development: Lessons for Innovation Policies
(November 1996: Eastern Germany)
40. R&D Subsidies at stake: In search of a rationale for public funding of industrial
R&D (18-19 April, 1996: Gent)
39. Re-Inventing the Future? Foresight at the edge of chaos November 1995:
Stockholm.
38. Technology, Innovation and (un) Employment 15-16 June, 1995: Rotterdam.
37. Research Co-operation with Countries in Transition (1-2 December 1994: Vienna).
36. Innovation - Applying New Ideas for Profit (25-26 May 1994: London).
35. Financing the Early Stage technology Company in the 1990's an international
perspective (8-9 November 1993: Montreal).
34. National Networks to Global Links (21-22 June 1993: Dublin).

33. User-producer relations in the innovation process (26-27 November 1992: Helsinki)
32. Lean Company: Model and Practice, Context and Policy (11-12 May 1992: Stuttgart).
31. Technical Competence and Firm Strategy - Implications for Public Policy (5-6 November 1991: Stockholm).
30. Innovation and Society: in Search of New Linkages (16-17 April 1991: Noordwijk aan Zee).
29. Technological Innovation and the Service Sector (23-24 October 1990: Vienna).
28. R&D Organizations in the 1990's (14-15 May 1990: London).
27. The Internationalization of R&D (16-17 October 1989: Ottawa).
26. The future of the Rural Economy (22-23 May 1989: Shannon).
25. Interfirm Innovation Dynamics (October 3-4 1988: Stuttgart).
24. Regional Innovation Supporting Services (December 7-8 1987: The Hague).
23. Investment in training (April 23-25 1987: Paris).
22. Home Information's (November 3-4 1986 London).
21. Engineering education for Industrial Development (June 9 - 10 June 1986: Stockholm).
20. Financing of Entrepreneurial Ventures and Innovation (30 September - 1 October 1985: Vienna).
19. International technology transfer: Promotion and Barriers (May 6-7 1985: Ottawa).
18. Regional Innovation Policies and Programmes (November 19-20 1984: Limerick).
17. Policies for industrial Innovations - Situations and Perspectives (May 7-8 1984: Bonn).
16. Technological Information from abroad: scientific attaches and other public institutions, industrial needs and public policies (December 1983: The Hague).
15. Technological culture and success in industry (6-7 June 1983: Paris).
14. Evaluating the effectiveness of government innovation policies (22-23 November 1982: London).
13. Industry university relations (10-11 May 1982: Stockholm).
12. A systematic approach to innovation (23-24 November 1981: The Hague).
11. Regional innovation policy: technology policies or regional policies (1-2 June 1981: Sophia Antipolis).
10. New entrepreneurship and the smaller innovative firm (9-10 June 1980: Limerick).
9. National innovation policies (1-2 December 1980: Berlin).
8. Trends in collective industrial research (26-27 November 1979: London).
7. Industrial Innovation and Government regulation (11-13 June 1979: The Hague).
6. Technical change and employment (13-14 November 1978: Paris).
5. Government procurement policies and industrial innovation (6-7 June 1978: Dublin).
4. Small and medium sized manufacturing firms: their role and problems in innovation, government policy in Europe, the USA, Canada, Japan and Israel

(21-22 November 1977: The Hague).

3. The current international economic climate and policies for technical innovation (November 1977).
2. Government aid to technical change in the mechanical engineering industries (19 April 1977: Karlsruhe)
1. Government direct financial assistance to industry, programmes, experiences and trends (12-13 October 1976: London)

◆ 最近のイノベーション政策関連の資料紹介として

R&D INVESTMENT IS KEY TO INNOVATION AND ECONOMIC GROWTH

英国 DTI (Department of Trade and Industry) が出版した「the Future and Innovation Unit of the UK's 2001 (<http://www.c3e.fr/Inteval/>) R&D Scoreboard」は、企業が潜在的な成長力を十全に発揮するためには研究開発投資が決定的に重要であることを明らかにした。

2.1.3 International Evaluation Research Group (INTEVAL)

<http://www.c3e.fr/Inteval/>

◆ 目的と活動の概要

政策・プログラム評価の国際的研究グループである INTEVAL は、1986 年以来、毎年、各国政府、会計検査機関、大学、民間セクターに属す評価専門家と会合を開催している。現在メンバーは 20 ヶ国以上にわたる。

国際的な視野に立って、電子メールやサブグループによる年次会合によって研究をすすめている。結果は、共著の書籍や Comparative Policy Analysis Series (Transaction Publishers) のシリーズとして出版される。

15 年間の活動によって、研究グループは評価の理論と実践に関してかなりの量の資料を収集している。

◆ 研究グループの会合歴

1986 年にブリュッセルで最初の会合を開いた。以後 9 年間はブリュッセルの International Institute of Administrative Sciences の支援を受けた。政策評価の理論と実践において国際的な比較研究をどう活用するかが焦点であった。

第 1 期のグループメンバーは、Dr. Ray C. Rist, Prof.dr. H.-U. Derlien, Germany, Prof.dr. R.V. Segsworth, Canada, Prof.dr. W. Jenkins and Prof. A. Gray, UK, Dr. E. Albaek and Dr. S. Winter, Denmark, B. Eriksen, Norway and Dr. G. Arvidson, Sweden plus the only founding mother Dr. M.L. Bemelmans-Vidéc. である。

以後の会合は次のとおりである。

- Brussels, 1986 hosted by IIAS
- Paris, 1987 hosted by the Ministry of Telecommunication and Post
- Leiden, 1988, hosted by the State University of Leyden
- Aarhus, 1989, hosted by the University of Aarhus

- Stockholm, 1990, hosted by the Swedish National Audit Office
- Bruges, 1991, hosted by the IIAS
- Ottawa, 1992, hosted by the Office of the Comptroller General, Canada
- Oslo, 1993, hosted by the Norwegian Government
- Lyon, 1994, hosted by the Centre for European Evaluation Expertise
- Madrid, 1996, hosted by the Institute of Fiscal Studies, Autonomous University of Madrid
- Seoul, 1995, hosted by the Hankuk University of Foreign Studies
- Washington, 1997, hosted by the World Bank
- Ireland, 1998, hosted by the Institute of Public Administration, Dublin
- Durham, 1999, hosted by the Durham University Business School
- Haifa, 2000, hosted by Haifa University.
- Stockholm, 2001 hosted by the Swedish Court of Auditors

◇ これまでの出版物

- COMPARATIVE POLICY ANALYSIS SERIES ,Edited by Ray C. Rist, Transaction Publishers.
- CARROTS, STICKS, AND SERMONS : Policy Instruments and Their Evaluation, Marie-Lousie Bemelmans-Videc, *Ray C. Rist, and Evert Vedung, Editors.*
- PROGRAM EVALUATION AND THE MANAGEMENT OF GOVERNMENT, *Ray C. Rist, Editor.*
- PUBLIC POLICY AND PROGRAM EVALUATION, *Evert Vedung.*
- MONITORING PERFORMANCE IN THE PUBLIC SECTOR, *John Mayne and Eduardo Zapico-Go, Editors . With a foreword by Joseph S. Wholey.*
- POLITICS AND PRACTICES OF INTERGOVERNMENTAL EVALUATION, *Jacques Tbulemonde and Olaf Rieper, Editors . With a foreword by Ray. C. Rist.*
- BUDGETING, AUDITING, EVALUATING : Functions and Intergration in Seven Governments, *Andrew Gray, Bill Jenkins, and Bob Segsworth, Editors . With a foreword by Ray. C. Rist.*
- CAN GOVERNMENTS LEARN?, *Frans. L. Leeuw, Ray C. Rist, and Richard Sonnichsen, Editors.*
- BUILDING EFFECTIVE EVALUATION CAPACITY : Lessons From Practice, *Richard Boyle, Donald Lemaire, Editors.*
- EL SEGUIMIENTO DE LOS RESULTADOS DE LA GESTION EN EL SECTOR PUBLICO, *J. Mayne, E. Zapico.*
- CARROTS, STICKS, AND SERMONS : Policy Instruments and Their Evaluation *Marie-Louise Bemelmans-Videc, Ray C. Rist, and Evert Vedung, Editors .*
- PROGRAM EVALUATION AND THE MANAGEMENT OF GOVERNMENT, *Ray C. Rist, Editor.*
- PUBLIC POLICY AND PROGRAM EVALUATION, *Evert Vedung.*
- MONITORING PERFORMANCE IN THE PUBLIC SECTOR, *John Mayne and Eduardo Zapico-Go, Editors . With a foreword by Joseph S. Wholey.*
- POLITICS AND PRACTICES OF INTERGOVERNMENTAL EVALUATION,

- Jacques Tbulemonde and Olaf Rieper, Editors. With a foreword by Ray C. Rist.*
- BUDGETING, AUDITING, EVALUATING : Functions and Intergration in Seven Governments, *Andrew Gray, Bill Jenkins, and Bob Segsworth, Editors.*
 - CAN GOVERNMENTS LEARN? *Frans L. Leeuw, Ray C. Rist, and Richard Sonnichsen, Editors.*
 - BUILDING EFFECTIVE EVALUATION CAPACITY : Lessons From Practice, *Richard Boyle, Donald Lemaire, Editors.*
 - Carrots, Sticks & Sermons; Policy Instruments & Their Evaluation Edited by *Marie-Louise Bemelmans-Videc, Ray C. Rist and Evert Vedung, Transaction Publishers, New Brunswick (U.S.A.) and London (U.K.)*
 - PROGRAM EVALUATION AND THE MANAGEMENT OF GOVERNMENT, *Ray C. Rist, Editor.*
 - The Politics and Practices of Inter-Governmental Evaluation, *Olaf Rieper and Jacques Tbulemonde, Editors.*
 - Budgeting, Auditing, and Evaluation - Functions and Integration In Seven Governments, *Andrew Gray, Bill Jenkins, Bob Segsworth, With a foreword by : Ray C. Rist.*
 - CAN GOVERNMENTS LEARN? Comparative Perspectives on Evaluation & Organization Learning, *Edited by Frans L. Leeuw, Ray C. Rist & Richard C. Sonnichsen.*
 - Building Effective Evaluation Capacity: Lessons From Practice, *Richard Boyle, Donald Lemaire, Editors.*

◇ 現在進行中の出版活動

次の4冊の作成が進行中である。

i) THE INTERNATIONAL EVALUATION ATLAS

Jan-Eric Furubo, Rolf Sandahl and Ray Rist, Editors.

- ・ 20カ国ならびに国際機関を取り上げ、そこでの評価システムを総覧している。

ii) EXPLORING EVALUATION AND COLLABORATIVE GOVERNMENT

Andrew Gray, Bill Jenkins, Frans Leeuw and John Mayne, Editors.

- ・ 政府がサービスを提供する際に、公的、民間ならびにボランティアな組織との共働が次第に求められるようになってきた。そうした流れの中で、政策とプログラムの評価はどう変化すべきなのかを探った。

iii) ASSESSING EVALUATIVE INFORMATION: prospects and pitfalls.

Reuven Schwartz, John Mayne, Jacques Tbulemonde, Editors.

- ・ 信頼性・妥当性の高い評価作業を実施しようとすればするほど、社会全体に‘査察’の目を張り巡らせ、息苦しさを生んでしまうという矛盾が生じる。この矛盾を解消できる評価のあり方を探る。

iv) EVALUATION UTILIZATION AND KNOWLEDGE MANAGEMENT : WHERE WE ARE WITH THE EVALUATION UTILIZATION DEBATE

Ray C. Rist, Marcus Spinatsch, & Nicoletta Stame, Editors.

<この先の出版予定>

- ASSESSING EVALUATIVE INFORMATION: prospects and pitfalls
Reuven Schwartz, John Mayne, Jacques Tbulemonde, Editors.
- EVALUATION UTILIZATION AND KNOWLEDGE MANAGEMENT (tentative title)
Ray Rist, Markus Spinatsch, Nicoletta Stame, Editors.

◆ INTEVAL のメンバー

- Erik Albaek
Department of Economics, Politics & Public Administration
Aalborg University (DENMARK)
- Goran Arvidsson
SNS Center for Business & Policy Studies (SWEDEN)
- Per Oyvind Bastoe
Managing Director, ErgoDialog AS (NORWAY)
- Jean-Claude Barbier
Director of Research (CNRS) Centre d'Etudes de l'Emploi "Le Descartes" (FRANCE)
- Marie Louise Bemelmans-Videc
(THE NETHERLANDS)
- Richard Boyle
Institute of Public Administration (IRELAND)
- Hans-Ulrich Derlien
Professor of Administrative Science, University of Bamberg (GERMANY)
- Stan Divorski
Assistant Director, Center for Evaluation Methods and Issues, US GAO (USA)
- Jan-Eric Furubo
The Swedish National Audit Bureau (SWEDEN)
- Andrew Gray
The Centre for Public Sector Management Research, Durham University Business School (UK)
- Katia Horber-Papazian
Professor, Institut des Hautes Etudes en Administration Publique (SWITZERLAND)
- William Jenkins
Darwin College, University of Kent, Canterbury (UK)
- Myoung-Soo Kim
Professor of Public Administration, Hankuk University of Foreign Studies
(REPUBLIC OF KOREA)
- Stefan Kuhlman
Fraunhofer Institute, Systemtechnik und Innovationsforschung (ISI) (GERMANY)
- Yoon-Shik Lee
Professor and Vice President for Academic Affairs, Department of Public Administration, Soongsil University (REPUBLIC OF KOREA)
- Frans L. Leeuw
Chief Review Officer, Open University and Utrecht University (THE NETHERLANDS)

- Donald Lemaire
Senior Regional Director, Quebec Regional Office, Department of Justice Canada
(CANADA)
- Jeremy Lousdale
National Audit Office (United Kingdom)
- John Mayne
Principal, Audit Operations, Office of the Auditor General (CANADA)
- Ragnhild Ovrelid
Directorate of Public Management (NORWAY)
- Dale H. Poel
School of Public Administration, Dalhousie University (CANADA)
- Olaf Rieper
Associate Professor, AKF - Institute of Local Government Studies-Denmark
(DENMARK)
- Ray Rist
c/o Operations Evaluation Department, The World Bank (UNITED STATES OF
AMERICA)
- Rolf Sandahl
The Swedish National Audit Bureau (SWEDEN)
- Robert (Reuven) Schwartz
Department of Political Science, University of Haifa (ISRAEL)
- Robert Segsworth
Department of Political Science, Laurentian University (CANADA)
- Nicoletta Stame
Universita di Roma "La Sapienza" (ITALY)
- Markus Spinatsch
Consulting for Public Policy and Administration (SWITZERLAND)
- Hilikka Summa
Directorate-General for Budgets (BELGIUM)
- Jacques Toulemonde
Center for European Evaluation Expertise (FRANCE)
- Marit Waerness
Directorate of Public Management (NORWAY)
- Eduardo Zapico-Goni
Deputy Director of Policy Analysis, Directorate of the Budget, Ministry of Finances
(SPAIN)

2.2 米国の行政内部の研修コース講師候補

◆ Department of Defense (DoD)

Office of Secretary of Defense (OSD)

▶ Leo Young

Title: Consultant to the Research Office in the Office of Secretary of Defense

◆ Defense Advanced Research Projects Agency (DARPA)

▶ Jane A. "Xan" Alexander

Title: Deputy Director

Address: DARPA

3701 North Fairfax Drive

Arlington, VA 22203-1714

Phone: 703-696-2400

Email: jalexander@darpa.mil

<http://www.darpa.mil/body/alexander.htm>

◆ Office of Naval Research (ONR)

▶ Dr. Ronald N. Kostoff

Address: ONR

800 N. Quincy St.

Arlington, VA 22217

Phone: 703-696-4198 Fax: 703-696-4274

Email: kostofr@onr.navy.mil

◆ Department of Energy (DoE)

▶ William J. Valdez

Title: Associate Director, Office of Planning & Analysis

Address: DoE Office of Science

1000 Independence Ave NW

Washington DC 20540

Phone: 202-586-4479

Email: bill.valdez@science.doe.gov

http://www.er.doe.gov/sub/leaders_org/Bio_Valdez.htm

◆ National Institute of Standards and Technology (NIST)

Office of Strategic Planning and Economic Analysis

(<http://www.nist.gov/director/planning/strategicplanning.htm>)

▶ Dr. Gregory C. Tasey

Address: NIST Program Office

Administration Bldg., Room A1000

100 Bureau Drive, MS 4710

Gaithersburg, MD 20899-4710

Phone: 301-975-2663

Email: gregory.tassey@nist.gov

http://www.nist.gov/director/planning/r&d_perf_meas.pdf

▶ Paul N. Doremus

Address: 同上

Email: paul.doremus@nist.gov

▶ Stephanie Shipp

Title: Director, Economic Assessment Office

Address: NIST Advanced Technology Program

100 Bureau Drive, MS 4710

Gaithersburg, MD 20899-4710

Phone: 301-975-8978

Email: stephanie.shipp@nist.gov

http://www.atp.nist.gov/eao/eao_main.htm

<http://www.atp.nist.gov/www/staff/sship.htm>

◆ National Institutes of Health (NIH)

▶ John Uzzell

Title: Director, Office of Evaluation, OSP, OD

Address: NIH

Suite 608, 6000 Executive Blvd., Rockville, MD, 20852, US

Phone: 301-496-6607 Fax: 301-480-9286

Email: uzzellj@odleml.od.nih.gov

◆ NASA

▶ Rich Beck

Title: Director, Resource Analysis Division

Address: NASA Headquarters

Washington, DC 20546-0001

Phone: 202-358-0000

参考資料 3. 評価人材が学習すべき重要文献

3.1 評価に関する主要総説文献

◆ Peer Review

1. Kostoff, Ronald N., "Research Program Peer Review: Principles, Practices, Protocols" (1997)

◆ 方法論

2. Kostoff, Ronald N., "Handbook of Research Impact Assessment. Edition 5. Summer 1995", NTIS (1995)
3. Bozeman, Barry (ed.), Melkers, Julia (ed.), "Evaluating R&D Impacts: Methods and Practice", Kluwer Academic Publishers (1993)

◆ Basic Research

4. Arnold, Erik, Balazs, Katalin, "Methods in The Evaluation of Publicly Funded Basic Research – A Review for OECD" Technopolis Ltd. (1998.3)
5. National Academy of Public Administration, "A Study of The National Science Foundation's Criteria for Project Selection", (2001.2)

◆ Socio-economic Impact

6. ETAN, "ETAN Working Paper, Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes", (1999.1)
7. PREST, AUEB, BETA, ISI, Joanneum Research, IE HAS, Wise Guys, "Assessing the Socio-economic Impacts of the Framework Programme", (2002.2)

◆ Program

8. Georghiou, Luke, Roessner, David, "Evaluating technology programs: tools and methods", *Research Policy*, Vol.29, 657-678 (2000)
9. ETAN, "ETAN Working Paper, Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes", (1999.1)
10. National Research Council, "The Advanced Technology Program: Assessing Outcomes", National Academy Press (2001)
11. Baranscomb, Lewis M. (ed.), Keller, James H. (ed.), "Investing in Innovation", The MIT Press (1999.6)
12. Becher, Gerhard (ed.), Kuhlmann, Stefan (ed.), "Evaluation of Technology Policy Programmes in Germany", Kluwer Academic Publishers (1995)
13. Ken Guy, "Strategic Options for the Evaluation of the R&D Programmes of the European Union", Technopolis Ltd. (1998.11)

◆ Policy

14.OECD, "Policy Evaluation in Innovation and Technology –Towards Best Practice", (1997)

◆ 総合/各国紹介

15. Callon, Michel, Laredo, Philippe, Mustar, Philippe, "The strategic management of research and technology", Editions Economica (1997)

16. OTA, "Research Funding as an Investment: Can We Measure the Returns? -A Technical Memorandum", NTIS (1986.4)

◆ 国際会議

17. Buhre, Susanne (ed.), Kuhlmann, Stefan (ed.), "Evaluation of Science and Technology in the New Europe", *Proceedings of International Conference*, Berlin (1999.7-8)

18. Philip Shapira (ed.), Kuhlmann, Stefan (ed.), "The 2000 US-EU Workshop on Learning from Science and Technology Policy Evaluation", Bad Herrenalb (2000.9)

19. European Commission Directorate General for Research, "The Contribution of Socio-economic Research to The Benchmarking of RTD Policies in Europe", Brussels (2001.3)

20. National Science Foundation, "Research Assessment: What's Next?", Washington. DC. (2001.5)

3.2 ピアレビューに関する重要総説文献の事例

評価の基本は当該専門領域の高い専門性をもつ科学技術ピアによる評価、ピア・レビューである。また、このピア・レビューを実際に運営し評価目的に沿った機能を発揮させるマネジメントが必要である。したがって、ピア・レビューには、評価人材の主要類型である、評価マネジメント人材と科学技術ピアが関連している。

高い質でなされるピア・レビューの基礎的諸原理に焦点をあてた重要総説論文に、Dr. Ronald N Kostoff の "Research Program Peer Review: Principles, Practices, Protocols" がある。これを以下に紹介する。

重要要素として、ここでは、

- 1) レビュー実施機関のシニア・マネジメントの高い品質のピア・レビューに対するコミットメントおよびピア・レビューを推奨する報償とインセンティブの配慮
- 2) レビュー・マネジャーのモチベーション
- 3) 評価者の能力と客観性
- 4) パネル間、分野間の正規化・標準化 (複数のプログラム/プロジェクト比較)
- 5) 評価基準の選択
- 6) 秘密性 (評価者の匿名性と被評価者の非匿名性)
- 7) コスト
- 8) 高い倫理基準の確保

があげられている。ピア・レビューに関連する人材の育成・活用に関して扱われる留意点でもある。

論文の構成は、

- 1.要約
- 2.導入、定義および背景
- 3.研究プログラムの定義（ピア・レビューの定義、背景）
- 4.ピア・レビューの原理（目的、質、マネジャーの役割、ピア・レビューアーの選定、評価基準の選定、秘密性：評価者と実行者の匿名性、ピア・レビューアーの客観性/偏向/公正性、パネルの標準化、反復可能性/信頼性、効果性/予測可能性、実行コスト、倫理の問題、ピア・レビューの代替的手法、二院制的評価、生産性に基づいた公式システム、サイエンス・コート、研究発展のための推奨事項）

からなる。下線部は人材の育成・活用にとくに関連する部分である。

RESEARCH PROGRAM PEER REVIEW: PRINCIPLES, PRACTICES, PROTOCOLS 研究プログラムのピア・レビュー：原理、実行、計画案

I. TITLE/ABSTRACT

タイトル/アブストラクト

研究プログラムのピア・レビューに関する原理、実行、計画案が述べられる。原理は基本的には包括的であり草案/プロポーザル/プログラムのピア・レビューと共に遂行機関のすべてのスペクトラムに及ぶピア・レビューに当てはまるものであるが、しかし、本論文の焦点は連邦政府機関における提案された研究プログラムおよび進行中の研究プログラムのピア・レビューである。

質の高いピア・レビューのための要因の完結したエグゼクティブ・サマリーの後に、1993年のGPRA（the Government Performance Results Act）の実行が連邦政府機関の研究プログラムのピア・レビューに及ぼした潜在的な重要性について述べる。次に、主要なピア・レビューの要素の長所と短所を説明し、そして以下の点について説明する；ピア・レビューの目的；ピア・レビューの質；ピア・レビュー・マネジャーの質に対する影響；ピア・レビューアーの選定；評価基準の選定；秘密保持（レビューアーと遂行者の匿名性）；ピア・レビューの目的/バイアス/公平性；ピア・レビュー・パネルの標準化；ピア・レビューの反復可能性/信頼性；ピア・レビューの効果性/予測可能性；ピア・レビュー実行のコスト；ピア・レビューの倫理的問題；ピア・レビューに代わるもの。

次に、連邦政府機関の異なったピア・レビューの実践と研究プログラムのピア・レビューを成功裏に行うための計画とプロセスの例が提示される。サイエンス・コートのようないくつかのピア・レビューの変種について説明され、そしてピア・レビューを改良するための研究上の要請について議論がされる。最後のセクションは、本文中の引用文献のみならず、さらに進んだ読者のための関連文献を載せた1500件以上の長い文献リストである。

II. EXECUTIVE SUMMARY - PEER REVIEW PRINCIPLES

エグゼクティブ・サマリー—ピア・レビューの原理

1993年のGPRA [GPRA, 1993]は、今年の秋に施行され、計画されたターゲットに向けての進捗程度を測るために、戦略的計画、年次遂行計画、達成指標を展開させていく事を連邦政府

に求めている。サイエンス誌のコンパニオン・ペーパー[Kostoff, 1997b]においては、GPRAが基礎研究に適用される場合にはピア・レビューを主要な手法として用いる事が推奨されている。しかし、GPRAのために研究プログラムのピア・レビューが効果的かつ効率的に用いられるためには、ピア・レビューは現状よりもっとよく理解され、改善され、標準化される必要がある。

本論文は高い品質のピア・レビューに必要な基礎的諸原理に焦点を当てている。本論文のターゲットは研究プログラムのピア・レビューであるが、ほとんどすべての原理は複合的なタイプのピア・レビューに適用可能である。ピア・レビューの文献を検討し、多くの実験的ピア・レビューを行い[e.g., Kostoff, 1988]、何百ものピア・レビューを運営した著者の経験をとおし、著者は高い品質のピア・レビューに大切な要素に関する以下の結論に導かれた[Kostoff, 1995, 1997a] (これは提案、プログラム、手続き、論文、研究者、学位論文のすべてに当てはまる)。

最も重要な要素はレビューを行う機関のシニア・マネジメントの、高い品質のピア・レビューに対するコミットメントと、それに関連してピア・レビューを推奨する報償とインセンティブの配置である。

第二に重要な要素は技術的に信頼性のあるピア・レビューを行うレビュー・マネージャーのモチベーションである。レビュー・マネージャーはパネル評価において質疑応答を先導し、彼らのコメントを要約し、勧告を作成する。いくつかの機関においては、レビュー・マネージャーはレビューのプロセスおよび評価基準を選択する力量を持ち、すべての機関において、恣意的でない方法で評価者を選ぶ能力がある。もしもレビュー・マネージャーが、意識的にせよ無意識的にせよ、評価者を選ぶのに最高の基準に従わなければ、レビューのプロセスが始まる前に、評価結果は大きく左右されてしまう。

第三に重要な要素は、評価者の能力と客観性である。各評価者は担当分野について専門的に能力があるべきであり、研究対象の複合的側面(当該研究分野; 隣接研究分野; 研究により潜在的に影響を受ける技術、システム、ミッション)を評価者グループ全体の能力の範囲が覆い尽くしているべきである。更に、評価グループの専門分野は評価される研究領域に限定されるべきではなく(ジョブが正しくなされたかという問題)、当該研究の最高レベルの目的がカバーする研究分野にまで拡張されるべきである(正しいジョブがなされたかという問題)。この拡張された範囲によって、評価グループは正しいジョブがなされたかという広域の問題を扱い、革命的に新規なパラダイムにも公平な考察を行える可能性が増大する。

複数のプログラムやプロジェクトの比較に用いられるピア・レビューに関して、第四に重要な要素は、パネル間および分野間の正規化・標準化である。ある程度の類似性を持つ分野間に関しては、広いバックグラウンドを持つ(パネル間の)共通の評価者を用いることによりある程度の標準化は達成できる[Kostoff, 1988, 1997a]。極めて離れた分野の間では、当該機関にとっての各分野の相対的戦略価値を配慮する必要があり、採点の厳しさや偏向性の補正が必要となる。極めて離れた分野の間でも、広いバックグラウンドを持つ共通の評価者を多様なプログラムやプロジェクトの評価に用いることで、ある程度の標準化は可能となる[Van den Beemt, 1997]。

第五に重要な要素は、評価基準の選択である[Delcomyn, 1991; Sutherland, 1993; Weinberg, 1989]。基礎研究提案の評価に際し、3つの主要基準は、研究メリット、研究手法、研究チームの質である[DOE, 1982; Kostoff, 1992]。ミッション指向の機関に支援された研究の場合、ミッションに係わる第4の基準が便利である。ミッション指向の基準がより基礎研究指向の提案をふるい落とさないためには、ミッションとの関連性を極めて広く解釈することが必要である。基礎研究の場合には、長期的基準よりも、(転換や有用性といった)短期的基準の

方が全体的な提案の質と点数との相関が高い[Kostoff, 1992]。第五の評価基準として研究の総合的質を選ぶことは重要であり、当該提案を評価するのに重要と評価者が考える評価リスト外の要素の影響を入れることを可能にする。例えば、評価者が当該提案は政府より産業界に支援されることがより適当であると感じたとしよう。この場合、この提案については、リストにある技術的基準の点数が高くても、総合点は低くなるかもしれない。

評価基準の第六の同様に重要な要素は秘密性である：評価者の匿名性および被評価者の非匿名性である[Altura, 1990; Clayson, 1995; Gresty, 1995; Neetens, 1995]。評価される研究の本質に関する正直で率直な見解が必要なら、評価者はレビュー・マネージャー以外には匿名でなければならない。提案（あるいは論文やプログラム）に対する否定的見解を表明する評価者にとって報酬はほとんどなく、評価者に対する懲罰や恨みは正直で率直な判断の科学に対する本質的利益をはるかに上回るだろう。

”ブラインド・レビューイング”つまり、被評価者の氏名と所属を評価者にわからないようにする事は、無名の研究者や評価の劣る研究所の研究者の仕事を公平に評価するという高貴な目的、あるいは性別といった個人的特性に基づいた偏向を除去するために用いられている[Ceci, 1984; Laband, 1994; Cox, 1993; Nylenna, 1994]。しかしながら、研究提案評価やモニタリング評価の研究によれば、プロジェクトの総合的質の決定に際し研究チームの質が最も重要な変数である事が示された[DOE, 1982]。被評価者に関する情報を隠して研究評価を行うことは、方程式を主要な項を除いてから解くことに等しい。研究者に関する重要な変数を取り除くより、評価グループの視点を広めて、研究ターゲットの’正しい仕事’の側面について言及できる追加的評価者を選ぶ事の方がより重要である。これにより、時流から取り残されたが論文生産が多く、被引用件数の多い研究が永続する事が無くなり、新しいパラダイムの新鮮な視点が公平に考慮されるであろう。

質の高いピア・レビューにとって第七の重要な要素はコストである[ASTEC, 1991; Buechner, 1974; Hensley, 1980; Kostoff, 1995, 1997a]。ピア・レビューの真の総コストは多額になりえる。しかしほとんどの報告事例では無視されるか過小評価される傾向にある。十分な数の専門家にパネルへの出席を求める質の高いピア・レビューの場合には、実際の総コストは直接コストをはるかに上回るであろう[Kostoff, 1995, 1997a]。総コストの中で大きな割合を占めるものは、スタッフや評価者、発表者を含む、評価の実施に関するすべての参加者の費消時間である。質の高い実施者と評価者にとって時間コストは高く、評価コストの総計は総プログラムコストの中で無視できない割合を占めることとなる。多くの提案が拒否され、種々の支援者への重複応募が通例となっているような状況においては、支援される提案1件当たりのピア・レビューのコストは大きく増加することとなる。提案（あるいは論文、プログラム）に関する質の高いピア・レビューのプロセスを設計するにあたってコストは無視されるべきではない。

最後の重要な点であり、多分、質の高いピア・レビューの基盤をなす要素であるものは高い倫理基準の維持である。真の専門家が評価者として望まれるとき、評価プロセス中には利害のバイアスや衝突が内在するので、科学的詐欺や科学的違法行為、秘密情報の流出、特権的情報へのアクセスにより不当な利益を取得するといった、倫理に関する多くの問題がある[Fielder, 1995; Goodstein, 1995; Gupta, 1996; Keown, 1996; Moran, 1992]。

どのような文書に評価者がサインしようと、最高の倫理基準を保つのにどのような望みを持つのであれ、彼らはアクセスする特権的情報に影響されざるを得ない。多くの経路を通じて知識移転は起こり、詳細な技術的説明を聞くことや詳細な提案書を読むことがおそらく最も効果的な移転方法であろう。倫理的な評価者は秘密を漏らしたり、ピア・レビューへの参加から不当な利益

を得るために意識的で明白な行動をとることはしないで、大きな研究事業を手助けし、種々のアイデアに触れることにより自分の思考が拡張することの満足を参加の報酬として受け取る。高い倫理基準を保つために、専門家共同体内の自己規制と、はなはだしい倫理的逸脱に対し法的手段に訴える事が精力的に行われなくてはならない。

III. INTRODUCTION, DEFINITIONS, AND BACKGROUND

導入、定義および背景

INTRODUCTION

導入

1993年、GPRA [GPRA, 1993]が施行された。GPRAは連邦支出のすべてのプログラムに適用され、3つのコンポーネントを持つ：戦略的計画、年次遂行計画および年次計画の達成度を示す指標がこれである。GPRAは法律となったので、この第3の要請、つまり達成指標を研究、中でも基礎研究の進捗と達成を適切に示すためにどのように実行するかを決定するため多くの連邦省庁間会合が開かれた。基礎研究の支援者と推進部門から出されたコンセンサスは、GPRAの法文が要求するところと健全な基礎研究に要請されるものとの間には大きなミスマッチが存在するというものである。

しかし、GPRAの法律では”もし、客観的、定量的、測定可能な形で特定のプログラム活動のパフォーマンス・ゴールを表すことが適切でない場合には、the Director of the Office of Management and Budgetは代替的形式をオーソライズしてもよい”となっている[GPRA, 1993]。Science誌のアーティクルにおいて[Kostoff, 1997b]、ピア・レビューは、文献計量的手法や他の手法で補足されつつ、基礎研究プログラムのGPRAに係わる健全性診断手法として用いられることが提案されている。ピア・レビューは、基礎研究プログラムの達成度測定に関し、GPRAの要求を満足させるためのより適切なツールであるというコンセンサスが研究者共同体において広がりつつある。もしGPRAの監督庁がこの考え方に同意すれば、連邦省庁間における研究プログラム・ピア・レビューの作業量は大幅に増加するであろう。

しかし、プログラム・ピア・レビューの作業量が変化するのみならず、必然的に評価の行為自体も変化するであろう。もし、GPRAが基本的に予算に関する道具であるなら[Brown, 1996]、業績に基づく予算編成プロセスへのインプットとなる業績評価結果の質は最高のものでなければならない。これらの業績評価結果を得るために選ばれた方法、つまり、プログラム・ピア・レビューと補足的な定量的手法は(プロセス選択および評価者選定などにおいて)より厳格で標準化された性格のものである必要がある。

本文書の目的は、研究プログラム・ピア・レビューを取り巻く基礎的問題と関心事について、関連する研究支援部門および監督、運営、推進諸部門の注意を喚起することにある。もしも、GPRAのフルスケールの実行以前にこれらの問題が包括的に表明されれば、業績に基づく予算編成プロセスを支援するのみならず、被評価研究プログラムに価値を付加できるようなピア・レビューを実行するための手続きが開発されるかもしれない。研究評価にかかわるより広いコミュニティにおける経験および発見事項を本文書が反映することを保証するために、論文ピア・レビューおよび提案ピア・レビューの文献における原理と発見事項が適宜活用される。これは、研究プログラム評価の問題に光を照射し、研究プログラム評価の文献中に存在するギャップを橋渡しするためである。

本論文には4つの主要なコンポーネントがある。テキストの本文(セクションII、III、IV)では研究プログラム・ピア・レビューを取り巻く基礎的問題が述べられる。セクションVでは、選

ばれた連邦政府機関における評価事例が要約される。セクションVIでは、連邦政府機関の最良の事例および本文中で採用された原理の多くを反映したピア・レビュー・プロセスの計画案が詳細に説明される。最後に、セクションVIIは、主要文献および関連するピア・レビューの参考文献の長い文献リストとなっている。最初に、ピア・レビューの問題を詳細に検討する段階に入るために、いくつかの定義と背景が提示される。

DEFINITIONS AND BACKGROUND

Research Program Definition

研究プログラムの定義

財政的には、研究プログラムは資金を供給された研究コンポーネントの集合である。これらの要素は、サブプログラムやプロジェクト、あるいは個々のワーク・ユニット（Principal Investigators-PIs）である。概念的には、生きている人間の体はそれを構成する分子の合計より大きいと同様に、プログラムはそのコンポーネントを合計したものより大きい。生きている人間の体が、分子を相互に、また、体のホメオスタシスの働きに結びつける知性を含むと同様に、プログラムは、そのコンポーネントを相互に、また、プログラムの総目的に結びつける知性あるいは内在論理を持つ。それゆえ、研究プログラムの内在的質は、コンポーネントとなるプロジェクトの質を単に合計したものではなく、プロジェクト相互間の構造的関係の質にも依存する。

研究プログラムの評価は2つの要素からなっていると見ることができる：1) ”研究のプログラムの評価(review of a program of research) ”、これはコンポーネントとなるプロジェクトの性質を検討し、一般的には綿密な技術的評価と呼ばれている、2) ”研究プログラムの評価(review of a research program) ”、これはプロジェクト間の構造的関係およびその外部環境の性質を検討し、一般的にはマネジメント評価と呼ばれている。これら二つの要素は一つの評価として作業的に統合することもできるし、別々に行う事もできる。

プログラムは、単一研究分野内で機関内あるいは機関間の場合；複合分野で機関内あるいは機関間の場合；複合分野で垂直統合された機関内あるいは機関間の場合；複合分野で複合機関かつ多国籍の場合；以上の他変種といった場合がある。本論文で述べられる名称としてのプログラムは機関内のものを前提しており、名称としての評価は機関内のものを前提としている。機関によっては分野別評価をおこなうところもあれば、多分野のマネジメント・ユニットごとの評価をおこなうところもある。また、ある機関においては、分野とマネジメント・ユニットが一致している。

Peer Review Definition

ピア・レビューの定義

ピアの伝統的な定義は”他者と同等の地位にある人” というものである。それゆえ、ピア・レビューは、ある人あるいは人々を同等の地位にある他の人々が評価することである。よって、本質的問題はいかに’同等の地位’が定義されるかである。

ジャーナルの研究論文であろうと、資金援助を得るための研究提案であれ、研究プロジェクトの達成度評価であれ、著者の知っているほとんどすべての研究ピア・レビューは被評価者や被評価グループの特定の研究分野の専門家をピア・レビューアーとして用いる傾向にある。評価者と被評価者の専門的知識の相対的レベルにより、評価者は事実上のピアであることもあればそうで

ないこともあるだろう。研究プログラムの評価の場合には、そのような専門家は、先程、“研究のプログラムの評価”として定義した深く技術的な部分については最も適している。この部分に関する焦点はプログラム内の研究プロジェクトの集合の本質、特に、その質と、達成度、現状の問題点および予期せぬ発見事項に置かれている。

先程、“研究プログラムの評価”として定義されたマネジメント評価の部分の焦点は、プログラム内の研究プロジェクト間の構造的関係に置かれている。この部分は、ミッションとの関連性と予算的妥当性、プログラムスタッフ、目的および手続きといった問題に係わるものである。この部分の問題を扱うためには、最初のピアグループに加えてさらにピアが必要となる。

本論文の目的上、研究プログラム・ピア・レビューの先程述べた二つの部分を扱うための要請をカバーするためには通常より広いピアの解釈が使用される。このようなピアの拡張された解釈は、先程述べたプログラム評価の二つの部分を単一のプロセスに統合する研究プログラム・ピア・レビューを実行する際に、著者が選定する傾向にあった評価者のタイプを説明するものである。このより包括的定義では、ピアとは、評価されている研究の特定の技術分野、評価されている研究に隣接する技術分野、評価されている研究が最終的に影響を及ぼす技術分野および評価されている研究が将来的に影響を及ぼすシステムとオペレーションの分野において専門家である人である。これら種々のタイプのピアには、評価されている特定の研究をはるかに越えて影響を及ぼし得る研究プログラムの種々の側面を検討することが求められている。

Research Program Peer Review Background

研究プログラム・ピア・レビューの背景

研究評価の方法は一般的に3つのグループに分けられる[Kostoff, 1995b, 1996a]: 定性的(例としてはピア・レビュー); 準定量的(例としては適及的手法); および定量的(例としては文献計量学的手法)。米国においては、世界の他の場所と同様、研究のピア・レビューが圧倒的に実際上選ばれている方法である[Salasin, 1980; Logsdon, 1985; Chubin, 1990; Chubin, 1994; Kostoff, 1995b; Stamps, 1997a; Wood, 1997]。現在、研究ピア・レビューが主に適用されているのは、使用頻度の大きいものから: ジャーナル投稿論文の評価; 提案評価; プロジェクトとプログラム評価; 教授陣のパフォーマンス評価; 学位論文評価である。

ピア・レビューの文献のほとんどは論文と提案の評価に焦点を当てている。例えば、1993年の文献調査[Speck, 1993]はピア・レビューに関する780の論文の要旨を集めた。そのうち、643の論文はジャーナルのピア・レビューに関するものであった。Armstrong [Armstrong, 1997]によれば、これらの101が経験的証拠を提供していた。プロジェクトやプログラムの基礎にある問題や原理について研究され、文献として公に報告されたものは比較的少ない。この結論は、SpeckとArmstrongの発見に補足される形で、著者が行った最近の文献サーベイにより非常にはっきりと確認されている。集められた文書の半分以上はジャーナルの編集者へのレターか論説(あるいはその等価物)であった。プログラム評価に関する論文は評価の技術的統計的結果に関する報告であって、ピア・レビューの基礎にある原理や問題にはほとんどあるいは全く重点が置かれていなかった。ピア・レビューに関するどの論文でも、原理は(主として)論文の評価あるいは提案評価に係わるのものであった。

研究プログラムのピア・レビューは、たとえ行われたとしても、論文や提案評価のように研究支援機関の間で整合的なものではない。プログラム評価は、非常にインフォーマルな個人的議論から、フォーマルな数十のパネル評価にまでわたるものである。プログラム評価を行う人のほとんどが評価を文書記録にとどめない。そして、ピア・レビューの文献の中でも原理や概念に係

わる論文のほとんどは研究プログラム・ピア・レビューをしたことが全くない人によって書かれている。結果として、研究プログラム・ピア・レビューの文献には二つの主要なギャップが存在する。第一に、量的に出版される論文が少なく、第二に、原理や概念に係わる論文のほとんどはプログラム評価を行う現実とほとんど関係を持たない。

これらのギャップのいくつかを識別し、扱うために、多くのピア・レビューに係わる問題が今検討されるであろう。これらの問題は、ピア・レビューや他の研究評価手法にまつわる問題の以前に行った評価からと、著者の最近のピア・レビュー文献サーベイによって生み出されたカテゴリ分類から選び出された[Kostoff, 1996a]。現セクションのすぐ後に続くテキストの本体で扱われているトピックの見出しは：ピア・レビューの目的；ピア・レビューの質；ピア・レビュー・マネージャーの質に対する影響；ピア・レビューアーの選定；評価基準の選定；秘密性（評価者と実施者の匿名性）；ピア・レビューの客観性／偏向／公平性；ピア・レビュー・パネルの標準化；ピア・レビューの反復可能性と信頼性；ピア・レビューの効果性と予測可能性；ピア・レビューの実行費用；ピア・レビューに係わる倫理的問題；ピア・レビューを代替するもの；ピア・レビューの更なる研究のための推奨事項である。

IV. PEER REVIEW PRINCIPLES

ピア・レビューの原理

OBJECTIVES/ PURPOSE OF PEER REVIEW

ピア・レビューの目的

ピア・レビューは多くの多様な目的をサポートする。ピア・レビューは資源を節約するクオリティ・フィルターの役を果たす：ピア・レビューのあるジャーナルはある最低限の品質以上であるはずであり、それによって、読者はこれらのジャーナルに掲載されている最高品質の文書に限られた時間を集中する事ができる；立ち上げや継続のためにピア・レビューにより選抜されたプロジェクトやプログラムは、ある最低限の品質以上であるはずであり、そして、貴重な労働力とハードウェア資源はこれらの選ばれた高品質の事業に集中される。ピア・レビューには、評価されている論文やプログラムに価値を付加し、質を改良する潜在力がある。ピア・レビューはプログラムに合法性と有効性のお墨付きを与える事ができ、その結果、プログラムへの注目度と援助を増大させる。ピア・レビューの諸目的には、効率的な資源配分機構から研究インパクトの信頼できる予測者までである。適切に行われた研究プログラム・ピア・レビューは研究支援者に、プログラムの質と、プログラムの関連性、マネジメントの質および方向の適切性の信頼できる指標を提供できる[Alassaf, 1996; Armstrong, 1997; Cram, 1992; Gabel, 1992; GERMANY, 1988; Kessler, 1992; Levine, 1988; Palli, 1993; Rainville, 1991; Ramsay, 1989; Stull, 1989; Wakefield, 1995; Wicks, 1992]。

ピア・レビューにより付加された価値を示す、いくつかの定量的研究の文献が存在する。例えば、近年の研究では、論文の質に対するピア・レビューと編集の効果 [Goodman, 1994]と、ピア・レビューと編集プロセスの独創的な論文を読みやすくする事への影響が評価された [Roberts, 1994]。彼らは、ピア・レビューと編集は、独創的な論文とその要旨を読みやすくすると共に、医学研究報告の質を改善すると結論した。彼らは、研究の質が改善されたかどうかには言及しなかったし、他の論文においてもその言及はない。

著者の経験では、研究プログラム・ピア・レビューのプロセスにおいて、三度価値が付加される時が存在する。第一には、評価の間の期間であり、この時には研究者は自分の仕事が質の高い

評価にさらされることを知りながら仕事をする。この実行期間において付加される価値は、専門家による評価が到来するという知識によって研究者がより高いレベルの仕事をするという事である。例えば、もし定期的に評価されることがわかっているなら、実施者は何十年も自分の課題に取り組むという傾向は減るだろう。プログラム・マネージャーは、もし評価が将来行われることを知っていれば、力のない実施者を力尽きさせるより、より継続的に要素プロジェクト間のバランスと関係を調整するだろう。

高速道路のよく知られたスピード・トラップの喩えがある。道路が厳重に警備されているという知識は、表示制限速度内に平均速度を押さえるのに十分である。だが、この地域では警官がスピードチケットを出すことが比較的少ないという事実は、スピード・トラップの効果の指標にはならない。この付加価値部分を実験的に検証できるかどうかを見るために、定期的に評価されるプログラムの研究の質と、たまに不定期に評価されるプログラムの質を比較する研究をすることは意義のあることであろう。

第二は評価の準備期間、特に発表を含む評価の予行演習の期間である。これはマネージャーにとっても研究者にとっても、非常に貴重な経験であり、それ自体で評価全体の費用と努力を正当と理由づけるものである。特に研究プログラム・ピア・レビューでは、準備期間は未解決の問題や優先事項に関する議論の活動の中心となり、質の良い発表を行えるために実質的な議論に火を付けることとなる。付加された価値は、表面的な発表の形の改善の中にはなく、プログラムの本質的価値の実質的増加の中に存在する。

第三は、実際の評価である。ここでは、独立した視点がパブリック・フォーラムに提出され、質の高い研究が再度肯定され、質の低い研究の運命として、強い忠告が出される。

評価結果によっては、価値が付加される四度目の時が想定される。もし評価結果が非常によければ、そしてプログラム資金が増加することになれば、少なくとも資金を受け取る者にとってまた、望ましくはより広い社会にとって、価値が付加されることになる。

最後に、いずれの評価プロセスも、研究の本質の現れではなく、研究の質に関するリアルタイムの判断を含むということを銘記するべきである。研究がその見込みを達成したかどうかを確かめるため研究の進化をたどるには、時間の経過が必要となる。ピア・レビューの判断が、研究の科学技術と社会への実際のインパクトにどれくらいよく相関しているかは、ピア・レビューの長期的価値の重要な指標となり、後出の「予測可能性」のセクションにおいてある程度扱われている。

ピア・レビューにより付加される潜在的価値の分類は次のように要約できる [Chubin, 1994] :

1. 研究資源配分の有効な機構 ;
2. 研究資源の有能な配分者 ;
3. 研究の説明責任 (science accountability) の主唱者 ;
4. 政策立案者が科学研究を指揮する機構 ;
5. 合理的な手順
6. 公平な手順
7. 科学的パフォーマンスの妥当でかつ信頼できる測定手段。

本論文の主要部の残りの多くの部分では、研究ピア・レビューにおいてこれらの望ましいゴールを達成する際の本質的・恣意的な障害物について検討を行う。潜在的バイアスや費用、現状の防御といったプログラム・ピア・レビューに関する多くの否定的側面が扱われる。本セクションでは諸文献で扱われなかったピア・レビューの潜在的に問題のあるもう一つの側面を手短に検討して結論とする。つまり、定期的に評価が行われるという知識は、非常に革新的であるが型破り

なアイデアの追求と発表を阻害するかどうかの問題である。実行者はこういったアイデアによって自分の信頼を損ないかねない状況で、パブリック・フォーラムでこれらのアイデアを表明することをためらうであろうか。換言すれば、ピア・レビューを行う事、特にパネルに基づいたプログラム・ピア・レビューは過激なアイデアの厳しい自己検閲をもたらしてしまっているのかという事である。定期的な評価プログラムにおいてアイデアが抑圧されるのかどうかを確かめ、もしも抑圧が存在するのならこの問題をどのように乗り越えるかを決定するためには、この分野における更なる研究が必要とされる。

QUALITY OF PEER REVIEW

ピア・レビューの質

ピア・レビューを実施する仕組みから、ピア・レビューの実例や、ピア・レビューの詳細な批判、およびピア・レビューの手順そのものに至るまで、ピア・レビューに関する多くの研究がこれまで文献で報告されている。以上において参考文献として挙げた検討や調査の中に含まれるピア・レビューと手順の資料に加えて、手順と批判の実例は次の文献中に見出される [Armstrong, 1997; Chubin, 1990; Chubin, 1994; Barker, 1992; Cicchetti, 1991; Cole, 1981; DOE, 1993; Frazier, 1987; Kostoff, 1996a; Wood, 1997]。

ピア・レビューの研究報告は手順の仕組みやとられる手続きおよび評価結果を説明するが、読者は評価での発見事項や勧告の質を確かめることができない。実際には、手続きや手順の質は、質の高いピア・レビューを作成するためのゆるやかな必要条件であり、しかし十分条件では決してない。様々な手順を用いて多くの有用なピア・レビューが実行された。そして、資料によって十分に裏付けられた現代的な手順(e.g., [DOE, 1993])は評価実行の効率性に貢献する一方で、高い質のためには手順以上のものが必要とされる。質の高いピア・レビューには多くの無形要素が入ってくる [Evans, 1990; Friedman, 1995; Goodman, 1994; Lundberg, 1991; Luukkonen-Grunow, 1990; McNutt, 1990; Vandenbroucke, 1994]、そして、そのうちで重要な要素のいくつかについては説明を加える。

本セクションでの基本的な前提は、研究の基本的業務のそれぞれにおいて内在する本質が存在するという事である。定義上、質の高いピア・レビューは、評価されている研究の本質の正確な見取り図を、本質が高かろうが低かろうが、提供するべきである。本質的な問題は、時間や長さといった第一義的測定のための物理学的基準とは異なり、研究の質を測定する絶対的基準は存在しない事である。現状では、研究の品質の評価は、評価者の視点や過去の経験に依存する主観的なプロセスである。これら不完全な条件下において、質の高い評価とは、次の二つの包括的条件が満たされた場合にできると定義される：1)高い能力を有する評価者の活用、2)偏向や衝突、詐欺あるいは不十分な作業の結果として、評価者の評価をさらに歪める事がない。

質の高いピア・レビューの手順には最低限、Ormalala [Ormalala, 1989]の研究から要約した次の条件が必要となる：

1. 特定の評価の状況に応じて、評価のための方法や機関および基準は選択され修正される必要がある；
2. 評価の異なるレベルでは、異なる評価方法が必要となる；
3. 評価研究を実施する際に、プログラムおよびプロジェクトのゴールには十分考慮が払われるべきである；
4. 評価の背景にある動機および評価と意思決定の関係についてはすべての関係者間でオープンに話し合われるべきである；

5. 評価目的は明確に定式化されるべきである；
6. 評価の信頼性は常に注意深く築き上げられるべきである；
7. 評価を設計するにあたっては、評価結果を効果的に活用するための前提条件を考慮に入れるべきである。

ピア・レビューの意思決定へのインパクトはピア・レビューの効果性の指標ではあるが、質の指標ではない。いいかげんなピア・レビューも理論上は決定に大きな影響を持ちえるし、うまく行われたピア・レビューが意思決定にきわめて小さい影響しか持たない事もありえる。ピア・レビューの質と効果性を分ける事が重要である。

ピア・レビューの質に関して当然生じてくる事柄として、著者の判断では名目上、研究プログラム・ピア・レビューの質に貢献する第一の要因ではないが、評価者による誤りがある。研究プログラム・ピア・レビュー者による誤りを検討した研究出版物を著者は知らない。最近の論文 [Armstrong, 1997] では、ピア・レビュー者によるジャーナル論文についての間違いや中身の無い仕事についての諸研究が説明されている。これらの結果から導かれる結論は、論文評価者による誤り産出の問題は取るに足らないという事である。Armstrong はジャーナル論文のピア・レビュー者は外部報酬を普通受け取らず、匿名であり、それ故にある場合には質の高い評価に必要な努力を行う動機がないかもしれないという点を強調している。

ジャーナル論文の著者は何百時間も研究を行って、論文出版の際には自分の名声を保つ必要があるのに対して、評価者のほうは、いいかげんな仕事をしてでも評判を落とす恐れは本質的にほとんどなく、相対的に少ない時間を評価に費やすので、この著者－評価者の間の協働関係には不均衡が存在する。法的体系においてはこの人間的弱点の存在が認識され、裁判官と陪審員が犯しえる誤りに対処する多層的ヒエラルキー上訴システムが確立されている。医療/法的体系においても、医療過誤システムにより確立された効果的上訴手続きが存在する。多分、科学専門職においても、論文の著者やピア・レビューにさらされる他の人たちの競技場を公平なものにし、最終的には正義が行われ、質が維持されることを保証するために、より公式の上訴システムの確立が必要とされている。最近の論文 [Stamps, 1997b] で、利害衝突の解決についての文献のレビューがなされ、科学分野のジャーナルにおける論文ピア・レビューでの紛争解決の手順（弁証的科学的摘要）を説明している。この手続きあるいは他の代替的手続きを修正すれば、他のタイプの科学的ピア・レビューへの適用が可能となる。

ほとんどの研究プログラム・ピア・レビューにおいて、匿名性と経済的動機づけの欠如に由来する弛緩した基準による技術的誤りは論文評価におけるほど深刻なものではない。少数のプログラム評価は、専門家による匿名のメールレビューによって行われるが（もしこれが適用される場合は、プログラムは各プロジェクトを個別に評価することによって評価されることになる）、プログラム評価の圧倒的多数は専門家パネルを用いて行われる。場合によっては、パネルメンバーはある程度の謝金を受け取るだろうが、とにかく、もはや匿名性は消失する。これらのパネルに参加することによって、彼らの名声が賭けられているのである。著者の経験では、パネルメンバーは偏見の明確な表明を抑制する傾向にあり、一般的には弁護する事の可能な事を述べる。この事が匿名のジャーナル論文評価に対し更なる保守主義を生むのかは、評価手順がどのように構成されているのかに依存しており、後の「秘密性」のセクションでより詳しく説明する。とにかく、研究プログラム・ピア・レビュー者による誤りの度合いに関する研究がなされるべきであり、もしもこれらのパネルが予算編成プロセスに実質的力を最終的に及ぼすとすれば、プログラム評価のためのある種の上訴システムが確立されなければならない。

IMPACT OF PEER REVIEW MANAGER ON QUALITY ピア・レビュー・マネージャーの質に対するインパクト

著者の立場からは、質の高い研究プログラム・ピア・レビューを生み出すのに最も重要な単一の要素は最高の質の客観的評価への機関のシニア・マネジメントの献身的努力であり、それに関連して、そのような評価を奨励するための報酬とインセンティブの配置である。質の高い評価を生み出すのに次に重要な要素は、技術的に信頼できる評価を行うために評価を管理している人のやる気である。この評価リーダーは評価プロセスを選択し、管理し、評価基準を選定し、評価者を選定し、パネル・レビューにおいて質疑応答を導き、メールあるいはパネル・レビューでの評価者のコメントをまとめ、当該プログラムを開始すべきか、継続すべきか、変更すべきかの勧告を行う。

もしも、意識的にせよ無意識的にせよ評価リーダーの偏向が、特に評価者選定プロセスにおいて表出すると、評価の方向は著しくそれによって影響を受ける。極端な事例では、評価結果は評価者の選定により、評価者が会合する前に、完全に決まってしまう。これは、プログラムあるいはプロジェクトのマネージャーや提案評価のマネージャー、ジャーナル論文評価担当の編集者について当てはまる。著者は、無作為のプロセスにより評価者が選ばれ、選定の際の偏向が排除されるタイプの評価においてはこのような事が起こった例を知らない。評価マネージャーによる意識的な評価者選定には潜在的かつ本質的偏向性が存在するので、評価者の評点と評価結果の間の数学的相関性[e.g., Cicchetti, 1991]は（これらの相関性は啓示的で洞察に富むものかもしれないが）常に点検されなければならない。

SELECTION OF PEER REVIEWERS

ピア・レビューアー（評価者）の選定

機関のトップ・マネジメントからの非常に強力な支援があり、公正で有能な評価リーダーの指揮があったとしても、評価の質は評価者の能力を上回ることにはできない。研究評価の際に考慮すべき能力の二つの次元は、当該分野における個々の評価者の技術的能力および評価グループの集団としての研究課題の種々の側面（他の研究インパクト、技術・ミッションへの考慮とインパクト、インフラ、政治・社会的インパクト）をカバーする能力である[Kostoff, 1995b, 1996a; Garson, 1980; Klahr, 1985; Marshall, 1996]。評価の質は評価者の偏向と軋轢により制限される。選定された評価者の偏向と軋轢をリーダーは知っているべきであるし、評価者同士でも知っているべきである。

パネル選定におけるよくある誤りの一つは、当該プログラムの要素技術分野の特定専門知識を持つ研究専門家に選択の幅を限定することである。これは、「仕事が正しくされているのかどうか」という問いには答えるが、「正しい仕事がなされているのか」という問いには答えていない。前者の質問は詳細な技術の質に係わり、一方、後者の質問はより広い意味における投資戦略のほうにより深くかかわる（投資戦略はプログラム・コンポーネント間の優先順位付けと資源配分の理論的根拠である）。後者の質問に答えるためには、プログラム全体の最高レベルの目的によってカバーされる分野の広い専門知識を持つ人たちも選ばれるべきである。彼らは、投資戦略をより客観的に扱う事ができ、要素技術分野の配分度合いや要素技術分野間の資源配分が妥当であるかどうかを判断できるであろう。これによって、評価グループは、正しい仕事がなされているのかという中心的な問いを扱う事ができるであろう。

論文であれ、提案あるいはプログラムであれ、ピア・レビューをめぐる主要な批判の一つは、それがオーソドックスで保守的なパラダイムを永続させ、現状の構造を脅かす新しいパラダイムを拒否する傾向にある事である。もしも研究プログラム・ピア・レビューの目的の一つが実際、技術革新が認められ、新しいパラダイムを持つ真に革新的研究が推奨され報われるという事であ

れば、仕事が正しくされているという問題を扱うための評価者と平行して、正しい仕事をするという問題を扱うための評価者を選定する事は、最高の重要度を持つ。

多くの現状の研究プログラム・ピア・レビューの最も深刻な欠点の一つは、パネル専門家が、仕事が正しくされているという問題に集中してしまっているという点であり、正しい仕事をするという問題を扱える専門家の事実上の不在である。この事は、著者が”The Pied Piper Effect”と命名したところの[Kostoff, 1996a]状況を招来させ得る。この現象は元々、ジャーナル論文の引用に関する特定の解釈を定義したものである。しかし、これはどのタイプのピア・レビュー（ジャーナル、提案、プログラム）の結論にも適用可能である。その最初の文献計量学的定義と次にプログラム・ピア・レビューへの応用を以下に述べる。

質とインパクトを計量する独立の指標として引用を用いることの主要な問題点の一つは、数字による結果が二通りに解釈できることであった。論文はそれが高い質を持つ事により、あるいは引用者が意見を異にする事が原因で、頻繁に引用され得る。しかしながら、最も油断のならない、さらに独立のモードで利用されている引用を排除する第三の解釈、即ち、”The Pied Piper Effect”が存在する。

研究のある分野において、現在主流となっている方法があると仮定しよう；例えば、癌治療のための化学／放射線／外科的方法（”The Pied Piper Effect”のより詳細な事例については[Kostoff, 1996a]を参照のこと）。以下の仮想的シナリオを仮定する：医療コミュニティの主流によって支持されていない代替的方法が存在する；50年後に癌の治療法が発見される；その治療法は今日の主流の方法と全く関係がなく、多分、今日の代替的方法の流れを引くものとする；医療コミュニティの主流によって是認されている今日の主流の方法は治療方法と全く無関係であるか、正反対でさえある。では、想像上、肯定的理由で頻繁に引用されている今日の癌に関する研究論文にどのような意味が付与され得るのであろうか。

この場合、その論文の高い引用度は、引用した論文の研究方向が正しいと研究者コミュニティを説得しようとする著者がしている程度の指標とはなるが、研究方法の本質的正しさの指標ではない。この事は、ミサイルを誤った標的に正確に飛ばす事と類似している。実際、高い引用度は、閉じられた研究者コミュニティ（引用される論文の著者と引用者達）の、より広い範囲のコミュニティ（これには政治家や他の資源配分者が含まれ得るだろう）に、自らの研究の方向が正しいという事を説得したいという熟慮された欲望を反映している。

これが、”Pied Piper”効果である。上記の医学の仮説的事例における、大量の引用件数はこの問題の程度や、正しい道から逸脱している程度を測る指標となるが、解決への進歩の指標とはならない。”Pied Piper”効果は、特に革新的研究の場合に、なぜ引用度や他の定量的指標が、研究のインパクトや質の信頼性のある評価において、幅広く構成されたピア・レビューの一部であり、従属的なものであるかの核心的理由となる。

”Pied Piper Effect”が研究プログラム・ピア・レビューの場合には何を意味するかは明白である。多くの専門家コミュニティは現状に満足しており、主流のオーソドックスな方法のために多くの個人的投資とインフラ投資を行っており、彼らの投資を時代遅れにする新たなパラダイムに脅威を感じている。もしも、ピア・レビューアーが、評価されている特定の研究方法の専門家だけで成っているならば、議論はその方法を用いる事がそもそも正しいのかどうかという事よりも（正しい仕事がなされているか）、その方法の詳細な部分が正しくなされているかどうかという問題をめぐるものとなるであろう（仕事が正しくなされているか）。そのような限定的評価の全体効果は主流の方法の継続に承認の判子を押すことであり（上で説明した高い引用度と類似している）、革新的思考への扉を閉じることである。

付録Iでは、今日用いられている最良の方法に近い、ピア・レビューアー選定方法を説明する。

それは純粋にランダムな選定手順ではないが、現状の選定での偏向の多くを取り除いてくれるし、ここで論じている大規模プログラム・ピア・レビューに適している。

SELECTION OF EVALUATION CRITERIA

評価基準の選定

研究評価基準は、機関が戦略的、政策的研究目的を定める道具である。評価者による基準へのきめの細かな応答は、ダウンストリームにおける意思決定のためのインプットとして貴重である。文書化された場合には、評価基準は、外部の集団にとって、戦略目的が実行されているという事の具体的な指標となる[Delcomyn, 1991; Eibeck, 1996; Kellie, 1991; Martin, 1981; Sutherland, 1993; Weinberg, 1964, 1989]。

個々の基準は数学的にベクトルの要素として見ることができる。ベクトル全体、あるいは評価の価値総点はその構成要素の点数の重み付け合計点として構成される。例えば、研究メリット (RM) とミッションとの関連性 (MR) という二つの基準が、研究プログラム評価で評価者により使われるように、評価機関により設定されたと仮定しよう。評価機関は各基準に同じ重みをつけるとしよう。では、制限をこれ以上つけないならば、最終的価値総点つまりプログラム全体の質 (OPQ) は $OPQ = 0.5 \cdot RM + 0.5 \cdot MR$ として計算される。

しかし、上記基準が評価者が重要と考える基準であることはほとんどないので、問題が発生する。上のケースでは、評価機関は重要と思われ、評価者に取り上げてもらいたい基準を二つだけ選び出した。評価機関は各基準につけられる重みと価値総点の算出計算方法も選んだ。各評価者は、研究評価にはどの基準が重要なのか、特定のプログラムではこれら基準の重み付けはどうか、これら基準からどのように最終的な価値総点を出すかという事について独自の意見をもっているのが衝突が起こる。何百もの種々のタイプのピア・レビューを扱った著者の経験では、評価を行う際に、評価者は研究パッケージ全体のゲシュタルトあるいは統合された見解を抱く。提供された要素基準は特定の分野において評価者の思考を刺激するのに役立ち、評価マネージャーにとって大切な問題を評価者が取り入れることを保証する。

先程の例では、機関の計算方法により得られる最終価値総点のベクトルと評価者の計算方法によるベクトルの間に重大な不釣合いが生じる可能性がある。二つのベクトルは十分異なり、一方が他方を代表するものとは全くなりえないかもしれない。例えば、機関は評価者に上述の計算方法を提供したとしよう。そして、研究メリットの定義 (当該問題の科学への重要性) は、研究方法 (当該問題を解くためにとられる方法) を含まないものとしてしよう。評価者が、評価されているプログラムについて、RM と MR の質は高いと思ったとしよう。しかし、評価者は、評価されているプログラムにおいて取られた研究方法は極めて悪いと重い、この特定の研究プログラムの総合的価値を決めるにおいて、研究方法は最も重要な基準であるとしてしよう。この場合、機関の基準と計算方法を用いることは、評価者が望んでいる結論と正反対の結果を出すことになる。たとえ機関が、評価者に基準の重みを自由につけてもよいという柔軟な姿勢を持っていたとしても、この例では、評価者の望む結論は、機関の計算方法を用い、恣意的に基準を重み付けした結果とも依然として正反対である。

著者が見つけたことは、熟達した評価者は通常、誠実な個人であり、上述のジレンマを解決するためにとられる方法は、原則の妥協よりも妥協の原則によってである。評価者は、評価される全研究パッケージの価値を直感的に判断する。そして、'リバーズ・エンジニア' によって、評価の計算方法が直感による総合結果と近くなるまで、(もし意識的でなければ) 潜在意識的に、基準の重みと評点を操作するのである。

これらの観察に基づき、著者は総合プロジェクト／プログラムの質の基準を取り入れることを推奨する（また、使用している）。この‘総決算’的評点は提示された全研究パッケージについての評価者の判断を明確にし、評価者が総合的な研究の質の重要な決定要因と感じるのに記述されていない基準（例えば、機関の妥当性）の影響を取り入れる事になる。この方法により、評価者の極めて深い信念に基づく点数に到達するために、‘リバース・エンジニアリング’を行う必要性が減じる。もしも、評価機関が依然として最終総合点に到達するのにその諸基準を用いたのであれば、評価者のベクトルと機関の計算方法によるベクトルを比較し、評価者の基準を無視する事による、評価者の視点での質のトレード・オフを識別する事ができる。

本論文で後出の、政府機関でのピア・レビュー実行のセクションで、研究プログラム評価基準の選定と重要性に関する、著者達による更に詳細な研究について説明する。一般的に言って、これらの研究によれば、評価者の最終的評点に対し最も影響がある基準は研究メリット、研究方法、および実行者の質である。更には、ミッションを持つ政府機関においては、ミッションとの関連性の基準が重要である。評価者の最終評点においては、技術への移転（あるいは有用性）といった、より短期的な関連事項の方が、支援者のダウンストリームのミッションとの長期的関連性よりも影響力があった。最後に、上述のように、単一の‘総決算’的基準を入れることは非常に重要である。

SECURITY: REVIEWER AND PERFORMER ANONYMITY

秘密性：評価者と実行者の匿名性

評価者の匿名性の問題については「ピア・レビューの質」のセクションで短い説明があり、結論としては、匿名性によっては評価者の評価の詳細な専門技術的質は改善されないという事であった。この否定的側面は、著者の視点では、評価者の匿名性のもたらす便益に比べれば、はるかに小さいものである。しかし、この結論については諸文献中で意見の一致があるわけではない [Altura, 1990; Berezin, 1994; Clayson, 1995; Debakey, 1990; Frei, 1993; Gresty, 1995; Knox, 1981; Neetens, 1995]。

ピア・レビューアーに本当に求められているものは、評価されている研究の本質に対する、可能などころでは厳正な技術的分析に支えられた、正直な視点である。評価者と被評価者を評価の際に、同席させることは（これは、論文評価、提案評価、プログラム評価に当てはまる；‘同席する’という語はそれぞれの場合に異なった解釈の仕方がなされなければならない）、専門的議論の詳細部分の質を研ぎ澄まし、「ピア・レビューの質」のセクションで説明した、諸研究 [Armstrong, 1997] において報告されている多くのタイプの誤りを除外する。

しかしながら、評価者と被評価者を評価の際に、同席させることは、多くの場合、研究の質に関する評価者の本音を表明する機会をなくしてしまう。研究論文や研究提案および、研究プログラムについて極めて否定的な意見を表明することに対する報酬はほとんどなく、その結果としての報復や恨みは、正直で率直な判断の表明のもたらす本質的便益をはるかに上回る。特に、研究プログラム・ピア・レビューでは、論文ピア・レビューよりも状況はより複雑である。プログラム評価においては、実質的に評価されているのは、プログラム・マネージャーと研究である。もしも、よくある事だが、研究プログラムのマネージャーの専門分野において、評価者が特別職レベル (bench-level) の専門家であり、更に、将来のある時点において、マネージャーの特定の研究プログラムに参加することに関心を持つかもしれないならば、率直だが否定的な評価は、そのプログラム・マネージャーから将来資金を獲得する自分の力に潜在的に深刻な結果をもたらすかも知れない。この場合、研究プログラム評価者として役に立つ真のピアを探し出すことは極めて

て困難であり、選定の手順において賢明な配慮が必要となる。

著者は完全に匿名の評価者の場合から被評価者や聴衆とずっと評価者が同席する場合までのすべての場合についてプログラム/提案評価を行った。著者の経験では、評価者の応答に種々の程度の率直さと正直さを生み出す、評価者の匿名性の階層レベルが存在する。

最も正直で率直な評価者の意見が聞けるのは電話による評価である。この場合、評価者は被評価者に対し完全に匿名である。このケースでは、評価者には研究についての情報（通常は文書で）が提供され、電話で口頭の応答が伝えられる。応答の率直さは、正しい仕事が行なわれているかを評価する点においてもっとも明白である。ここでは、全研究方法の完全さが俎上に載せられる。主要な方向性や、基盤的変化を危険にさらしているわけでは別にないので、評価者は、仕事が正しくなされているのかを批評する際にはより積極的になる。そして、被評価者の弁明はそれほど騒々しくない。

階層において次にくるのは、評価者が被評価者に対し完全に匿名である、文書による評価である。幾人かの評価者は文書でコメントを提出するよう言われる場合、コメントの率直さを手加減する傾向にある。しかし、評価マネージャーが匿名性を守る事を評価者が信用すれば、文書でも彼らはとても率直である。

匿名性の次のレベルでは、評価者と被評価者は研究内容の発表の場に同席するが、評価者は非公開のセッションで研究についての口頭および文書での評価をおこない、評価は無記名である。非公開セッションでの匿名性でさえ、非常に率直な議論と心からの意見交換をもたらす。

最後のレベルは、匿名性の不在であり、評価者と被評価者はすべてのプロセスにおいて同席し、すべての口頭および文書によるコメントは記名による。このタイプの評価は全く評価をしないよりはましだと言われているが、著者の経験では、この方法は専門家のピア・レビューの提供できる可能性を十分有効利用していない。

秘密性のコインのもう一つの面は、被評価者の名前と所属を評価者から隠す事である。この手順は”ブラインド・レビューイング”と名づけられている” [Blank, 1991; Ceci, 1984; Cox, 1993; Evans, 1990; Fisher, 1994; Johnson, 1995; Laband, 1994; McNutt, 1990; Nylenna, 1994; Rosenblatt, 1980; Shaughnessy, 1988; Sly, 1990]。この目的は、無名の研究者や、評価の低い研究所の研究者による研究を公正に評価する事 [Armstrong, 1997]、あるいは、考えられるところでは、ジェンダー等の個人的特長に基づいた偏向を除外するためである。ブラインド・レビューイング（およびその派生系としての”ダブル・ブラインド”評価、つまり評価者と被評価者がお互い匿名である評価）はおそらく論文評価に最も適している。ジャーナル論文のブラインド・レビューイングに関する研究のいくつかが報告されている [Fletcher and Fletcher, 1997; Fisher, 1994; Laband, 1994]。編集者は、ブラインド・レビューによる評価者の評価は高い質を持つと判断している；ブラインド・レビューによる評価者は以前より多くの論文を持つ著者により良い点をつけ、ブラインド・ピア・レビューを用いたジャーナルに載った論文は、ブラインド・ピア・レビューを用いないジャーナルに載った論文よりずっと頻繁に引用されている。

残念ながら、被評価者の身元情報を、評価されている研究から除外する事は、主要な項を除いた後で方程式を解くのに似ている。DOE の、Office of Basic Energy Sciences' 研究プログラム [DOE, 1982] に関するピア・レビューの研究（これは、おそらくプロジェクトの質のコンポーネントの統計的標本を用いた、研究プログラムの質に関する古典的研究である）の結論によれば、プロジェクト全体の質を決定する最も重要な変数はチームの質である。この事や他の類似の結果に基づけば、被評価者の身元情報なしに提案を評価する事は、誤った結果を生む可能性があるかもしれない。提案された良い研究トピックは、多数存在する。質の高い研究者は良い研

究トピックを扱った記録を持つのみならず、辛抱強さと鑑識眼のある思考により解決に向かって大きく進歩する。研究者が資金援助提案書を用意するのを手伝う多くのコンサルティング会社が、今日では存在する。これらのコンサルタントは素人をうならせるのに適当な専門語や政治的に正しい言葉遣いを熟知し、どのタイプの形式と提案組織構造が多くの決定者に訴えるものがあるかを良く知っている。研究者抜きにそのような提案を判断することは結局、形式を内容に優先させてしまうことになる。

とにかく、ブラインド・レビューの研究プログラム評価への適用可能性は極めて小さい。ほとんどの場合、パネル評価が用いられ、被評価者の身元情報を保護するには、特別な注意が必要となるだろう。チームの質に関する基準を使えないのと相まってプログラム・ピア・レビューでこの手順を用いる理由はほとんど無いように見える。諸文献中には、プログラム評価に関連してこのトピックを論じたものは存在しないようである。

OBJECTIVITY/ BIAS/ FAIRNESS OF PEER REVIEW

ピア・レビューの客観性／偏向／公正性

おそらく、すべてのタイプのピア・レビューで最も批判される側面は、評価者の最終勧告において、偏向の持つ役割であり、それが公平性に及ぼしていく影響である。ピア・レビューは文書の中でも口頭でも、性別や、人種、制度、地理、年齢に対する偏向、特に既成分野を守ろうとする「オールド・ボーイ」のネットワークへの保守的偏向を持つと非難されてきた。多くの研究努力が偏向と公平性のこの問題に集中された[Armstrong, 1982, 1997; Bailar, 1991; Daniel, 1993; Ehlen, 1996; Ernst, 1994; Ramasarma, 1995; Spitzer, 1994]。Armstrong[Armstrong, 1997]は、近年の大規模な研究[Speck, 1993]において、ジャーナル評価に関する実証的論文の半分近くはこの問題を扱っていると指摘している。

見出されたことは雑多である。最近の研究[Gilbert, 1994]では、JAMAによって受理された論文が、ピア・レビューの手順における参加者の性別に関連して、ピア・レビューと論文処理の点で異なる扱いを受けるか、受理される割合が異なるかを評価した。この研究の結論では、JAMAにおいては、編集者と評価者の特性において性別による違いが存在するが、ピア・レビューの手順や受理される事において明白な影響は存在しなかった。

他の研究[Peters, 1982]によれば、評価者は、無名あるいは名声の劣る研究所の著者に対し差別するという事が見出された。NSFの提案評価が違うパネル[Cole, 1981]によって評価しなおされた研究においては、研究所の評判や勤務年数、アカデミックな肩書き、地理的位置および他の変数が取り入れられた。その結論では、NSFによるピア・レビュー・システムには本質的にシステムの偏向は存在しなかった。DOEのOffice of Basic Energy Scienceの研究[DOE, 1982]によれば、研究所のプロジェクトおよび研究所以外のプロジェクトに関する結論は評価者による偏向で歪められてはいなかった。

1992年の報告書の連邦による投資を組織する共通のフレームワークに関するガイドラインを解説したセクションでは、偏向および衝突について詳細に述べられている[NAS, 1992]。その原則6（プログラム評価）には次のような文章がある：“政府研究開発プログラムを評価する現在の努力は、場合によっては、議会やエグゼクティブ・ブランチへの年次報告書が、ミッション機関の被雇用者が、自分が評価しているプロジェクトが継続することに直接的利害を持ちつつ、作成されたという事実により妨げられてきた。研究開発プログラムの専門的評価および前商用段階にある研究開発プログラムの国家経済福祉への貢献度の専門的評価は、プログラム・管理や資金援助の決定に直接的役割を持たない非政府系のグループによって行われるべきである。”

偏向／公平性の問題の基礎的パラダイムは、すべての評価者は同等に扱われるべきであるという事である；すべてのプレイヤーにとって、競技場は公平でなければならない。残念ながら、この高邁な哲学を実行するにあたって、政治的正しさのルールが科学的根拠のルールに優先する。目に見える更なる公平性への動機付けが、おそらく秘密性に関する前セクションで扱われた、'ブラインド・レビューイング'のようなピア・レビューの概念を生み出す主要なドライバーとなっている。”ブラインド・レビューイング”への下降は、結論として、過去の記録（チームの質）という核心的評価基準を除外することであり、それは評価手順の質を劣化させることになる。

しかしながら、後の代替案のセクションで幾人かの研究者によって提案されているように、過去の記録に圧倒的は重要性を付加する事は、機能的バランスを、正しい仕事が行なわれているかという側面とは反対の仕事が正しくなされているかという研究側面を強調することにつながり、多くの面で両刃の剣である。これは、ほとんど過去の業績はないが困難な研究課題を解決する非常に良いアイデアを持ち、これらの課題を扱う非常な能力があるかもしれない若い研究者にとっては重大な障害となり、「オールド・ボーイ」のネットワークと現状を維持する事になる可能性がある。前セクションでの”Pied Piper Effect”の説明で示されたように、これは非常に深刻な結果をもたらすかも知れない。このパラドックスの解決策は、研究者の身元情報を除外する事ではなく、パネルの視野が拡大するように評価者の選定を行う事である。研究ターゲットに関して正しい仕事が行なわれているかの点を扱うことができ、時代遅れだが被引用件数の多い研究が永続的に行われる事がなく、専門知識のプールが常に補充される事を保証できるパネリストを選べということである。

NORMALIZATION OF PEER REVIEW PANELS

ピア・レビュー・パネルの標準化

ピア・レビューは単一の研究のみにも使え、多くの種々のタイプの研究を比較するためにも使える、診断的な手順である。比較の目的に用いられる場合には、核心的問題は、分野間の比較およびパネル間の比較に意味があるように、種々の専門分野を評価する異なったパネルの評価結果を標準化する事の問題である。例えば、評価されている種々のタイプの研究の本質の相違を、種々のパネルの偏向や、基準のパネルによる種々の解釈と、基準を適用する際のパネリストの厳しさの程度の違いから分離することは、これらのすべての要因を含む評点とコメントのみが示されたときに、どのようにして可能であろうか。この標準化の問題はおそらくピア・レビューの最も難しい側面であり、文献計量学のような研究評価の他の側面にも標準化の困難性の問題は当てはまるのである[Braun, 1982; Kostoff, 1997c; Schubert, 1996]。

分野間のピア・レビューを検討した研究のほとんどは、主要な分野のカテゴリーごとに別々に結果を提示している[e.g., DOE, 1982; Cicchetti, 1991; Cole, 1981]。彼らは本質的に巧みに問題を回避している。機関のミッションに対する分野の重要度により分野が選択され、保持されるという戦略的見地から研究が眺められたときには、このカテゴリーによる分離は正当なものであるが、この分野ごとの分離は、質の比較指標としてのピア・レビューの価値を著しく減少させる。文献計量学のような定量的評価方法では、異なった分野間の参照基準が開発され、研究を順位付けするための計量手順が作成された[Schubert, 1996]。これにより、広い包括的意味においては、分野間の相対ランキングの比較ができるのであるが、しかし、ある分野（例えば音響学）で定義された参照基準が、分野内のプログラム比較（例えば、水圏音響学 vs 航空音響学）に使えるのかという問題が生じる[Kostoff, 1996a]。

著者は、先にあげた多くの変数の存在のために、完全に満足できるピア・レビューの標準化方

法を見た事がない。しかしながら、研究提案を評価するためにオランダの STW では一つの興味深い標準化方法が使われている [Van den Beemt, 1991, 1997]。質の等級付けではないが、専門的コメントが専門家ピアにより提供される。このコメントと提案者の応答が、20 の異なった提案について、様々な分野からの 12 人に渡される。この 12 人の '陪審員' は、独立したメール・レビューで評点をつける。本質的に、標準化は、すべての提案を 12 人の共通の陪審員が担当することによりなされる。

著者は、パネル間の標準化をいくらか改善するための二つの方法を用いた事がある。最初の方法は、すべてのパネルで共通の個人を幾人か用いることである。1980 年代後半 [Kostoff, 1988] の新規に促進された研究プログラムの一連の競合において、著者はすべての種々の分野のパネルの議長を務めた。この結果、種々のパネル間で標準化が小規模に行われた。更に多くの個人をすべてのパネルで共通に用いることは標準化の度合いを更に進めるであろう。そしてこの意味で、評価の際にシニア・マネジメントが同席することは更に標準化を進めるであろう。明らかに、トピックとしてパネルがより緊密に関連していればいる程、種々のパネルに共通の個人がする専門的な貢献をより貴重なものとなる。

第二に、主要な分野（例えば、物理学と生命科学）の合計平均点は次の二つの要素によると仮定された：提案されたプログラムの本質の相違および評価者の採点の厳しさの度合の相違。標準化するために、主要分野の総平均点の差異の一部が除かれた。これにより、点数のつけ方の厳しさの差異が取り除かれたと考えられる。試行錯誤の結果、50%の補正ファクターにより、すべての評価に出席した聴衆にとって、直感的に妥当に見える結果が得られた。この標準化の手続きには、機関にとって戦略的価値のある分野の適切な反映を保存し保証するという付加的便益があった。

標準化を行うこの方法には第二番目の解釈の仕方がある。もしも、研究が戦略的部分と質的部分を持っていると見ることができ、評価者の評点は質的部分のみを表しているとするならば、補正は戦略的部分を組み入れて修正しているものと見れるだろう。例えば、「ライフ・サイエンス」パネルのプログラムに対する平均点が 5 点とする。「エンジニアリング・サイエンス」パネルの平均点は 10 点とする。両分野は機関にとって同等の戦略的価値をもつものと仮定する。戦略的価値は評価者の評点（これは、ミッションとの関連性を含むプログラム全体の質の評点とする）と同等の重要性を持つ。ゆえに、標準化された総合点は $FOM = 0.5 * STRAT + 0.5 * score$ と計算される。これにより、二つのパネルの評点の差異は 5 点から 2.5 点に減少する。最終的に標準化された評点に到達するために、この補正ファクターは分野内の各プログラムの粗点に用いる事ができる。

もし、ピア・レビューが GPRA を支援するために用いられるものとするれば、信頼性のためにある種の標準化の手続きが必要となる。標準化のための既存のスキームが、特に離れた分野間では、非常に限られた妥当性しかない事を考慮すると、これは困難である。もしも、GPRA が研究予算に影響を及ぼすために用いられるのであれば、評点を標準化する妥当な手続きが本質的に欠かせないのであるが、現在それは存在しない。ピア・レビューの研究にとってこの領域は非常に肥沃である。

REPEATABILITY/ RELIABILITY OF PEER REVIEW

ピア・レビューの反復可能性/信頼性

物理学的システムの実験において、結果の信頼性を測るために尋ねられる主要な質問の一つは、結果の反復可能性に係わる。別の実験室において、同じコントロールされた条件下で、同じ実験

をすれば、同じ結果あるいはそれについての妥当な複製が生み出され得るのだろうか。ピア・レビューにおける類似の問題は、言葉を替えて、信頼性 (reliability)、反復可能性 (repeatability)、一貫性 (consistency)、一定不変性 (uniformity) などと名づけられ、諸文献中で極めて注目されてきた[Baillar, 1991; Ceci, 1982; Cicchetti, 1976, 1979, 1991; Cole, 1991; Colman, 1991; Crothers, 1993; Daniel, 1993; Gorman, 1991; Halpin, 1986; Kiesler, 1991; Kraemer, 1991; Laming, 1991; Luce, 1993; Marsh, 1989; Roediger, 1991; Rosenthal, 1990, 1991, Rubin, 1992]。

物理学のシステムにおいては、残念ながらピア・レビューに常に持ち込まれているわけではない、二つの推論的概念が存在する。これらは正確さ (precision) と精度 (accuracy) の概念である。正確さは測定値が複製される正確さを表し、精度はある絶対的価値や基準と測定値の関係を表す。

何百もの引用を用いた、論文と提案グラントのピア・レビューの信頼性に関する包括的研究 [Cicchetti, 1991] において、信頼性は、一般的に種々の尺度で定義された：内的整合性、レフェリー間の同意 (レフェリー間の同意の程度)、時間的安定性。これらの定義における信頼性は、上で定義した正確さの類似物であるように見え、ここで、精度の問題は定義に入っていないようである。この研究によれば、最も共通の尺度は、ある時点のレフェリー間の同意である。この研究の本質的結論は、検討された様々な科学領域において：1) 論文およびグラントの応募において、質が劣っていると認識されるものに関する同意の程度の方が、質が高いと認識されるものに関する同意の程度より大きい；2) 科学的調査の良く定義された (つまり、特定化され、専門化された) 分野においては、よく定義されていない (つまり、一般的であり、焦点のぼやけた) 科学的関心の分野よりも、受理される割合がより高く、用いる評価者の数が少ない；3) 偶然性を補正した後のレフェリー間の同意の程度はかなり低い。

しかしながら、この研究の注釈や参照された研究の説明においては、評価者の真にランダムな選定の問題は扱われておらず、それゆえに、この研究の結果の意義には検討の余地がある。例えば、これらの条件において、高い信頼性とは何を意味するのか。これは、評価者が論文/提案の本質を正確に識別し報告できるという事を意味するのかもしれないし、トピックにたいする極端な (肯定的あるいは否定的) 偏向性を持つ事が理由で評価者は選ばれ、評価マネージャーは同様の偏向性を持つ評価者を選ぶことに優れた手腕を発揮したという事を意味するのかもしれない。

更には、慧眼なマネージャーは全く異なった視点と専門知識を持つ評価者を選ぶだろうから、偶然性の要素を取り除いた後のレフェリー間の同意の程度は実際低くあるべきであると考えられる学派が存在し、この結果として、問題の種々の側面に評価者は敏感であるべきであるというのである。この視点からは、あまりにも同意の程度が高いことは弱点のある証拠であり、マネージャーが広い見聞に基づく決定を行うに必要な意見の全スペクトルをそのシステムは実現していないという事である。

全米科学財団 (NSF) から資金が出され、評価者の二つの集団を用いた、NSF の提案の研究 [Cole, 1981] によると、反転率 (一方のグループの決定が他方のグループの決定により覆される割合) はおよそ 25% である。完全にランダムな手順によれば、反転率は 50% になるだろうから、結論として、あるグラント応募の運命は、約半分は提案および主任研究官の性格により決定され、約半分は明らかにランダムな要素によって決定されるという事になる。また、観察される評価者間の意見の不一致の大部分は、おそらく、良い科学とはどのようなものであるか、あるいはどのようなものであるべきかという事に関する専門家の間の実際の正当な意見の違いの結果であるという結論が出された。

同様に、研究プログラム評価の信頼性に関する研究は、おそらく、特にパネル評価については、

そのような研究に必要な複製のための費用と努力のため、また、同一の手順がはたして本当に複製されるのかどうかという疑問が原因となって、諸文献の中には存在しないようである。進行中の研究プログラムおよび提案された研究プログラムに関する著者の経験では(その経験のほんの一部については文書にし、数学的分析を行ったのだが[Kostoff, 1992])、実際的な目的にとっては信頼性は十分なものである。[Kostoff, 1996a]においてより詳細に説明されているように、ピア・レビューは、進行中の研究プログラムおよび提案された研究プログラムについて、それが優れているかあるいは劣っているかについて合意を形成する事が可能であるが、一方、そのどちらでもないはるかに広い中間的な領域をカバーする種々のプログラムに関する意見が存在する。この中間領域に属するプログラムに関しては、その運命は、選定される評価者により、より敏感に左右されることになる。もしもピア・レビューの基本的目的が、非常に優れたプログラムに資金が提供され、プログラムが継続される事、および劣ったプログラムが中止されるか大幅修正される事にあるのならば、ピア・レビューという道具の持つ能力は、その要求によく一致している。

著者のプログラム・ピア・レビューの信頼性に関する経験は、以上で述べた事柄や類似の諸研究よりも、いくらか更に肯定的であるように思われる。なぜか。それは多分、どのようにピア・レビューが実行されるかという事に、おおいに原因がある。諸文献中で報告されている多くの提案および論文評価においては、評価者間および評価者と著者/提案者の間のやりとりが極めて少ない傾向がある。おそらく、良くて反証の提出が一回あるくらいである。この独立性には疑いなく価値があるし、すべてのプレイヤーを一堂に会して質疑応答を行うよりもはるかに安上がりである。

著者の行うピア・レビューでは、評価者と発表者の間に、広範囲にわたる質疑応答がある。評点が出される前の専門的情報の交換の間に、多くの誤解や解釈の違いが解きほぐされる。最初の評点は評価者が独立して採点する。次に、評点の違いが議論され、評価者には自分の評点を修正する機会が与えられる。通常、最終的な評点における差異は縮まる。著者の観察によれば、この評点の分散の減少は、より強力なあるいは騒々しい論者の支配が原因ではなく、提示された資料の本質を、各評価者がよりよく理解する事によるものである。したがって、ジャーナル論文分析[Chicchetti, 1991]で用いられたように、評価者間の同意の程度を信頼性の尺度とするよりは、研究プログラム・ピア・レビューにとっては、信頼性のより良い尺度とは、パネルにおいて、上で示したように、交流がなされた後の平均的パネル評点における一致の度合いであろう。

EFFECTIVENESS/ PREDICTABILITY OF PEER REVIEW

ピア・レビューの効果性/予測可能性

ピア・レビューの予測可能性の問題は、技術予測の信頼性に直接的に影響を持つ。研究のピア・レビューを実施する機関にとっては、評価者の評点を、機関のミッションに及ぼすダウンストリームのインパクトに関連づけることが望ましいであろう[Abrams, 1991; Van den Beemt, 1991, 1997]。要素評価基準に基づく評価者の評点を研究提案やプロジェクトの評価結果と関連させようとする、少数の研究がこれまでになされた(e.g., [DOE, 1982; Kostoff, 1992])。研究論文に関する評価者の格付けを、その論文がある期間にわたって引用される件数と比較するという研究もいくつか行われた[Bornstein, 1991a; Bornstein, 1991b]。論文の質とインパクトに関する評価者の評価とある期間にわたってその論文が引用される件数との相関係数は比較的低い。Bornstein は、ピア・レビューの信頼性と妥当性に関する膨大な調査の後で、次のように結論付けている：“ピア・レビューの手順の信頼性および妥当性と同程度の信頼性と

妥当性を持つ評価ツールを含む研究を出版しようとするなら、疑いなく、この精神測定学的に欠陥のあるこの手法を含む研究は受理されないと考えられるだろう” [Bornstein, 1991b]。

著者の知る限りでは、提案のピア・レビューによる評点/評価を、提案研究が技術とシステム、オペレーションに及ぼすダウンストリームの影響に関連付けた大規模な研究については、単独の形であれ共著の形であれ、存在しないが、この目的へ向けてのある程度の努力は開始された [Van den Beemt, 1991]。このタイプの研究は、長期間にわたる詳細なデータ追跡システムを必要とするだろう。よって、(研究それ自身のためではなく) 機関のミッションへの研究のインパクトを評価する予測的ツールとしてのピア・レビューの価値は、確固たる文書化された根拠に基づくというより、信念に依拠している。

COSTS OF PERFORMING A PEER REVIEW

ピア・レビュー実行のコスト

ピア・レビューに関するもう一つの問題はコストである [ASTEC, 1991; Buechner, 1974; Hensley, 1980; Kostoff, 1995b, 1996a]。ピア・レビューの真の総コストは、これから示すように、かなりの額となる事があるが、無視されたり、報告されてもほとんどの場合過小評価されがちである。ピア・レビューには多くの異なったタイプがあるので、一般的なピア・レビューに対する総コストの概算方法を示すことは極めて難しい。それにもかかわらず、ピア・レビューの総コストに関するおおよその桁を示す概算値の以下の例を考えてみよう [Kostoff, 1996a]。

仮に或る、中間評価ピア・レビューには、一つの研究所で年・プログラム当たり100万ドルを必要とする、と想定されたい。ピア・レビューの運用方法として、専門家の一団に2日間研究所の所在地への出頭を求め、主要な調査担当者からの説明を聴取することとする。評価団(パネル)は、研究、技術、研究運営等の10名の専門家から構成されると仮定し、8名の主席担当研究員がパネルにプロジェクトの内容を提示すると仮定する。各パネル・メンバーに負荷されるコスト(月給プラス時間外手当)は年当たり15万ドルと仮定する。また、各主席担当研究員に負荷されるコストは年当たり12万5千ドルと仮定する。会食費や旅費のような直接費は、6千ドルから8千ドルのあたりになるであろう。何らかの礼金などがあればこのコストは更に増える。

評価者、説明者、スタッフおよび評価の陪席者などが評価に費やす時間の総計などの間接費は、12万5千ドルの範囲にあり、またこれには少なくとも以下のものが含まれるであろう。

1. 評価に先立って評価者に読んでもらうための背景資料の作成、パネルにおける説明資料の作成、経営幹部への予行演習の実施等に使う説明者の時間。(概算4万ドル; 80人日)。
2. 背景資料(文書、報告書、計画書)を読み、評価のために旅行し、会合で時間を費やし、報告書を書く、等のためのパネル・メンバーの時間(概算4万8千ドルから6万ドル; 80-100人日)。
3. 評価者を選びかつ評価を依頼し、評価会議を設置し、研究現場との調整を行い、報告書を作成する、等に使う研究機関スタッフの時間(概算1万ドル; 20人日)。
4. 評価会議における陪席者(研究現場の経営幹部、他の現場職員、他の研究機関代表者、等)の時間(概算2万ドル; 40人日)。

以上の検討の主な結論は次のようになる。すなわち、パネル型のピア・レビューを本格的にやろうとして、十分な数の専門家にパネルへの出席を求めるとき、実際の総コストは直接コストを遥かに越すであろう。この結論は郵送タイプのピア・レビューにも当てはまるであろう。郵送タイ

ブのピア・レビューの総コストは、旅費が要らないために、パネル型のピア・レビューの総コストより少ないとは言うものの、郵送タイプのピア・レビューにとって、直接費に対する総コストの割合は極めて高いものとなるであろう。いずれのタイプの評価にとっても、総コストの中で大きな割合を占めるものは、評価の実施に関与するすべての参加者の費消時間である。地位の高い評価会議推進者と評価者が参加するとき、時間コストは高くなり、評価コストの総計は総プログラムコストの中で無視できない割合を占めることとなる。特に設備集中型より、人員集中型のプログラムに対してはそうなる。

ETHICAL ISSUES IN PEER REVIEW

ピア・レビューにおける倫理の問題

研究という職業においては、科学上の詐欺や、科学上の違法行為、秘密情報の漏洩、特権的情報へのアクセスによる不正な利益の取得を含む多くの倫理的問題がある。これらの領域においては、法的小よび文書化されない／語られない同意や罰則が存在する。ピア・レビューの従属的目的の一つには、論文レベル[Fox, 1994]であれ、提案、プログラムレベルであれ、特に詐欺と違法行為に関して、高い倫理基準を維持するという事がある。多くの詐欺や違法行為は文書化された専門的産物において発生してきたので、この分野のピア・レビューへの適用報告例のほとんどは、ジャーナルのピア・レビューからのものである[Fielder, 1995; Goodstein, 1995; Gupta, 1996; Keown, 1996; Mokrasch, 1988; Moran, 1992; Southgate, 1992]。これらの分野での倫理基準の維持は、研究者共同体による自警による傾向がある。著者は、プログラム・ピア・レビューにおいて、詐欺や違法行為が発見されたのを見た事がない。また、諸文献中においてもそのようなケースを読んだ事がない。

あらゆる形の研究ピア・レビューの根底にある、基本的な倫理のパラドックスが存在する。評価の手順に信頼性を持たせるためには、正しい仕事が行なわれているのかという事や仕事が行なわれているのかという事に関して、専門家を用いなければならない。大衆的な意見とは反対に、過去に行われた実験や評価実施時の個人的観察に基づく著者の経験によれば、どの研究分野においても、真の専門家は極めて少数である。Armstrong[Armstrong, 1997]は類似の結論を論文ピア・レビューについて述べている。つまり、評価者は当該論文の著者と近接した分野で研究しているが同一の問題について研究しているわけではないので、その結果、著者よりも評価者のほうがその問題全体についての経験が少ないというのである。よって、パネルに真の専門家に入ってもらうため、少なくとも研究が正しくなされているかという側面を評価するためには、比較的小さな研究者コミュニティに接触しなければならないのである。通常、このコミュニティのメンバーはお互いに顔見知りであり、更には共同研究者であったり、研究上の競争者であったりする。彼らは、資金や賞、名声、昇進、あるいは他の認知を求めて競争するのである。だから、真の専門家が評価者として欲しいときには、内在的に偏向性／利害衝突がプロセス中に存在するのだ。

普通、研究プログラム・ピア・レビューでは、評価される研究の秘密を守らねばならないという、評価者がサインする文書が存在する（或いはすべきである）。しかし、実際上は、所有権に係わる微妙な情報を承認無しに使用する事を制限するものは、文書化されず語られる事のない倫理基準の厳守なのである。法的な保護も存在するし、最近では、自分の秘密や財産権に関わる研究が、評価者が個人的な利得として結果を不法に取得する事によって、侵害を受けたと感じた人によって、法廷に事件として持ち込まれることもある。

どのような文書に評価者がサインしようと、至高の倫理基準を厳守するべくどのような希望を

もっているとしても、彼らは自分がアクセスする特権的情報によって影響を受けざるを得ない。知識の移転は多くの経路を通して起こり、詳細な専門的説明を聞いたり、専門的提案書を読むことはおそらく、二つの最も効率的な経路であろう。専門的資料へのアクセスによってもたらされる倫理的ジレンマの効果的解決法は、原則の妥協ではなく、妥協の原則である。倫理的な評価者は、秘密を漏らしたり、ピア・レビューへの参加により不正な利益を得るといったような意識的で公然の行動をとることはせず、むしろ参加の報奨として大規模研究事業を支援し、種々のアイデアに自分の思考プロセスをさらした事による満足を受け止めるのである。研究プログラム・ピア・レビューをより大規模に用いる事が現実化すれば、また、評価結果を予算の決定に反映させるために用いるようになれば、ここで述べた倫理的基準のいくつかの厳守を保証するための努力が更に必要となってくる。

ALTERNATIVES TO PEER REVIEW

ピア・レビューの代替的手法

本論文は、ピア・レビューの使用にかかわる多くの問題を識別してきた。種々の問題の実行の際の深刻さは各適用対象によって異なるのではあるが、これらの諸問題は、プログラム評価や提案評価と論文評価への種々のピア・レビューの適用全体に概念的に共通したものである。ピア・レビューの修正版および全くの代替的手法については、ピア・レビューの最も問題のある側面を克服するために試みとして、多くの提案が出されている[Forsdyke, 1991; Greene, 1991; Roy, 1981, 1984, 1985; Smith, 1988; Wick, 1996; Wood, 1997]。これらの代替案の考えのほとんどは(他の形のピア・レビューにもこれらのアイデアの要素部分は適用できるが)、特に研究提案ピア・レビューに集中している。広く知られている代替案のうち二つをここで説明し、批判を行う。

Bicameral Review

二院制の評価

幾人かのカナダの科学者によって[Berezin, 1995; Forsdyke, 1991]、プロジェクト選定のためのピア・レビューの修正版が近年提案された。この手法は”二院制評価(Bicameral Review)”と発案者のForsdyke博士によって名づけられ、その要点は次のようなものである。

二院制評価の構造は、基礎研究の成果を予測する事に内在する不確実性のために、研究資金提供システムは非常に誤りやすいという仮定に基づいている。もしも、評価システムが非常に間違いを犯しやすいなら、システム設計においてその間違いやすさを考慮に入れておかなければならない。不確実な条件下での意思決定に関する二つの原則は：1) ある程度の客観性を持って評価できそうなパラメーターに重きをおく、2) 資金を分散投資して危険を防ぐ。

二院制評価においては、グラント応募や大多数の遡及的部分(提案者の過去の記録)と少数の将来的部分(提案された研究)に、別々のルートに分けられる。遡及的部分のみがピア・レビューを受ける。将来的部分は、予算の妥当性の側面のみについて、機関の内部評価を受ける。ピアはただ生産性を評価するだけでなく、受け取るドル当たりの生産性を評価する事が要求される。さらには、応募者の経験を計算に入れなければならない。若い研究者は、履歴を確立するまで、より多くの資金の”ロープ”を与えられる(不確実性の余得)。資金配分は、既存の資金を出すか出さないかのくっきりとした線引き方法の代わりにスライディング・スケールで行われる。ファンディング・スケールの極めて上位のもののみが、妥当な期間内に研究を完了するのに必要な

資金のすべてを獲得するだろう。プロジェクトの格付けがファンディング・スケール上で下がるにつれて、要求額に対し提供される額の割合は減少する。

Productivity-Based Formula Systems

生産性に基づいた公式のシステム

過去の成功は将来の業績の最も良い指標であり、少数の集団を継続的に妥当な期間の間支援していくことは成功の確率とシステムの効率を高め、ほとんどの革新的科学は最小限のマイクロ・マネジメントによってなされるという原則に基づいて、ピア・レビュー以外の代替的手法が提案された[Roy, 1981, 1985]。この代替的手法は、研究者が本質的に過去の業績に基づいて資金の提供を受けるべきであるという事を提案し、資金配分のための計算法を提供する。一つの計算法において[Roy, 1985]、提供される金額は、出版件数と高等学位の数、ミッション機関からの研究援助額および産業界からの研究援助額の重み付けした合計に比例する。そして、資金は研究ユニット(部など)へ出される。繰り返しになるが、基本的原理は、将来性よりも過去の業績の方が公的説明責任にとりかはるかにしっかりとした土台を提供するという事である。研究ユニットに新たに加わる研究者については、基本配分公式に追加的に値が加えられる事になる。

Author's Commentary on Alternatives

代替的手法に関する著者の注解

理想的には、研究提案評価は、アイデアの源泉が何であるかにかかわらず、最も可能性の大きいアイデアに資金を配分できるべきである。そのような手順は、優れた業績を持つ評価の確立した研究者や、業績は劣るが評価の確立した研究者と業績がない新たな研究者からのアイデアを含むべきである。それは、学会および政府や産業界の研究者を含む事ができるべきであり、また、一人での計画から非常に大きな機関の計画まで、また、研究結果を報告することに関し種々の立場や文化をもった古典的研究や非古典的研究を含む事ができるべきである。最終的決定に到達するのに資金配分の手順は、基礎研究の結果を予測することの不確実性を認識しながら、最良の専門的判断を取り入れるべきである。

ここで取り上げた二つの代替的手法は優れた業績を持つ評価の確立した研究者への資金提供を非常に強調する。これらの手法は、業績をどのように決定するかという点において異なる。つまり、二院制の評価ではピアを用い、生産性に基づいた公式による評価では公式を用いる。両方法とも、提案された研究の将来性の評価について真の専門家を用いる事を最小限にしている。実際に実施する際には、これらの代替的手法は既存のピア・レビューの手順と、その説明を一読して想像する程にはそれほど大きく変わらないだろう。以前述べたように、分析によると、実施者の業績の婉曲語法である「チームの質」が、実施中の研究と提案された研究について総合評価点を決定する際の支配的要因である。よって、既存の手法と代替的手法は両方とも実際には業績記録に非常に重点を置いているのである。代替的手法と既存の手法の間の真の違いは、著者の見解では、提案の将来性の部分を評価する際の専門家の使い方にある。

二つの代替的手法は提案提出に係わる費用をある程度減少させ、評価者の偏向の影響を減じ、競争者の新奇なアイデアを剽窃する機会を大きく減らし、評価手順にかかる不必要な時間消費を削減するが、しかし、いくつかの欠点を持つ。

提案された概念に関する専門家の判断を除外して極端に業績を強調することは、研究の舞台へ

の新規参入に対する障害を増加させる事により、オーソドックスな主流の方法を永続化させることになる。提案された研究を判断するための専門的知識の欠如は、選定手順において非専門的要素が支配的になる事につながり、評価において形式が内容に相対的に優先する事になる。

ゼロサム・ゲームにおいては、二院制の評価手順は、スライディングスケールの存在と足切の除外により、'最良'の提案から'最悪'の提案へ資金のいくらかを移転するように見える。しかし、これによりすべて或いはほとんどすべての研究者に資金を配分することにより'セーフティ・ネット'を提供する事になる。

生産性に基づいたシステム手法はサイエンスのコンパニオン論文[Kostoff, 1997b]で扱われている現在のGPRAの手法に幾分類似しており、似たような多くの欠点を持つ。研究を評価するための独立した手法として、どの指標や指標の組み合わせを用いる事でも誤りに陥ってしまう。選択された指標は研究の質を表す妥当な指標かもしれないがそうでないかもしれない;その指標の信頼性を確認するためにはピアの解釈が必要とされるのだ。公式に基づいた手法は、研究者達を基本的理解よりも数字に表れるアウトプットの達成に駆り立てるというネガティブな可能性を持つ。

生産性に基づく手法は、方程式の循環的システムに類似しており、初期条件が間違っていれば、最終評点も間違いになるであろう。例えば、方程式の項の一つがミッション機関から研究がもらう金額だとしよう。その研究チームは立法措置により'指定された'大きなグラントを受け取っていたとしよう。この事により更に、大学院の学生数と生産された論文数がより良い数字になるかもしれない。そして、研究プログラムの本質とは必ずしも関係していない高い総合評価点が出てしまうかもしれない。この誤った初期条件に基づいた資金配分は毎年繰り返され、その'指定'措置が終了した後でさえ続く、永続的システムとなるだろう。それなので、もしも何らかの公式や定量的指標の組み合わせが使用されるのならば、上述のような状況が起こるのを回避するために、専門家のピア・レビューがそれに同伴し、それを従属的なものにしなければならない。

これらの代替的手法および他の類似した性質の手法は、ピア・レビューによる選定手順では最良の研究を生み出せず、そのような不正確な手順を続ける事を正当化するには提案を作成するのにあまりに膨大な時間とエネルギーがかかるという前提に立っている。この基本的前提には疑問の余地があり得る。ピア・レビューには欠点や限界がある一方で、最良の研究者や最良のアイデアが資金無しでやっていける証拠はまずないし、上記の代替的手法が状況を改善する証拠は更に少ないのである。

SCIENCE COURT サイエンス・コート

概念評価のための非標準的ピア・レビュー法として、サイエンス・コートがある。法的手続きにおいてと同様に、それには明確に定義された弁護人と批判者、陪審員などがある。これは、ユニークで力強さを秘めた手法であるが、道具と同じで、もし理解されずうまく適用されないならば、誤用される可能性がある。著者によって1977年に、磁気核融合オフィスにおいて、核融合概念に関する代替案の評価に適用された[DOE, 1978]。

評価のために選択された一般的な形式は、選定された評価者による、弁護の手続きを含むパネル・レビューであった。手順における主要な要素グループは、運営委員会と評価パネル、弁護団、批判者達であった。これらの参加者と評価におけるその役割については以下で説明される。

運営委員会は核融合オフィスの代表者から構成されていた。この委員会の主要な責任は(1) 評価を運営する、(2) 評価基準を定義する、(3) 評価パネルのメンバーを選定する、(4) 評

価に際し評価パネルを支援する、(5) 評価者の結論と勧告を受けて、核融合オフィスへの最終報告書を作成するという事であった。

評価パネルは、プラズマ物理学者と核融合反応システムの専門家およびユーティリティ産業の代表者から構成されていた。このパネルには、検討される概念のどれかを積極的に推進する人間は含まれていなかった。もしあるパネルメンバーにとって間接的に利害衝突が起こってしまう場合には、そのメンバーはその特定の概念の検討の場には参加しなかった。パネルはすべての概念の専門的評価に対し責任があった。

概念の弁護団はその特定の概念について研究を行っている科学者やエンジニアであった。弁護団は、反応炉具体化の技術や魅力と科学的結果や予測について情報提供と弁護をおこなう責任があった。主任弁護人が弁護団の活動を指揮するために指名された。

批判者たちには、特定の概念にとって重要な物理や工学の分野における特別な専門知識を持っている者が選ばれた。批判者たちの責任は重要な物理的質問や技術的質問を探し出す事であり、実験結果と理論的モデルの評価において評価パネルを手助けする事であった。ある概念の提唱者が場合によっては、他の概念の評価において批判者の役を演じる事があった。主任批判者が批判者達の活動を指揮するために指名された。

参加者のだれもが(弁護団、批判者達、評価パネル)外部の専門家を必要に応じて用いる事が許された。この手続きでは、著者が今までに見た中で最も多く議論が交わされ、重要な問題点が最も多く浮かび上がった。しかし、これは標準的なパネル評価に比べて時間のかかるものであった。

RECOMMENDATIONS FOR FURTHER RESEARCH IN PEER REVIEW

ピア・レビューの更なる研究のための推奨事項

以上で述べられた問題点や関心は特に研究プログラム・ピア・レビュー、また他の形のピア・レビューを実行する際の、多くの溝や欠陥を照らし出している。最も重要な推奨事項は、論文ピア・レビューと提案ピア・レビューにおいて分析された側面について、研究プログラム・ピア・レビューにおいても研究を始めるべきだという事である。プログラム・ピア・レビューの実行や原理についての研究は非常に数少ない。また、もしプログラム・ピア・レビューがGPRAを支援するために拡張されるのであれば、効果的で信頼できる比較診断装置にするために、その強みと弱みをはるかによく理解する事が必要となる。

すべてのタイプのピア・レビューを巡る中心的問題点の一つは予測の信頼性を欠いているという事である。研究提案と実施中の研究プログラムのピアによる評価をその研究の将来インパクトに関連付けるためには更なる研究が必要となる。現状では、種々の予測モデルが有効であるとするデータは存在しない。研究の産物の進化を、様々な変形段階において辿ることができるデータベースが必要となる。そのようなデータベースを持つ事で、ピア・レビューの予測モデルのみならず、文献計量学的モデルおよび他の定量的予測モデルの有効性をも確認する事ができる。このデータベースにより、多くの種々のインパクトに関する予測信頼性を決定できる。これには、関心研究分野へのインパクトと隣接研究分野へのインパクト、技術へのインパクト、システムへのインパクト、オペレーションへのインパクトなどが含まれる。

ピア・レビューの結果の有効性と信頼性についての議論はCicchetti[Cicchetti, 1991]の論文とDanielの論文[Daniel, 1993]およびCicchetti論文が載ったジャーナルの号の他の論評中に見出される。有効性と信頼性を改善するためには、評価者の人数の最適値に関する研究; 著者の匿名性が結果に影響するのかを確認し; ピア・レビューを行うために人を訓練する

ことは信頼性と有効性と共に評価の質を改善するのかどうかを確認する研究が必要となる。

種々のタイプのピアのグループ分けおよびピア・レビューの産出物の質に関する比較研究は非常に少ない。次の諸点に関し研究がなされるべきである：メール・レビュー vs. パネル・レビュー、英国モデル（著名な研究者の代わりに評価の職業的専門家を用いるピア・レビュー） vs. 標準モデル、パネルの規模、評価者の専門性のタイプ、プロセスにおいて評価者と被評価者によって費やされる時間、また、これらの変数と産出物の質の相関を見る事。結果の中心をなすのは、評価コストが産出物の質とどう相関するかという点と、評価コストが種々の変数によってどのように影響されるかという点であろう。

多くの変数(分野、パネルなど)における標準化の問題は以前に未知のものとして識別された。様々な研究プログラム・ピア・レビューの変数に関しどのように標準化を行うかについて研究が行われるべきであるという事は繰り返し述べる価値がある。

一つのピア・レビューのシナリオに関しピア・レビューの時間と費用の合計の大雑把な推計を本論文では行ったが、種々のピア・レビューのシナリオを比較する際にはより正確な時間と費用の推計が必要とされるだろう。多くのタイプのピア・レビューが存在するので、膨大なデータ収集が必要となるだろう。しかし、ピア・レビューの総費用はかなりの額にのぼり、首尾一貫した質を保ちながら費用削減を行う事は、種々のタイプの提唱された研究のゴールの一つであろうと考えられるので、経済的見地から見て、膨大なデータ収集とピア・レビューの費用推計の改良された手続きの開発は正当な事であろう。

提案研究評価とプログラム評価へのエキスパート・システムと知識ベース・システムの適用はピア・レビューを補完するであろう。この方法に沿っての研究は数少ない、しかし近年の学位論文 [Odeyale, 1993] とそれに続く研究 [Odeyale, 1994a, 1994b] がこの問題を詳細に扱っている。これらの先進技術の適用をピア・レビューを補完するものとして正当化するにははるかに多くの研究が必要となるだろう。しかし、この方法に沿った更なる研究は真の利益が存在する可能性を決定するだろう。

ピア・レビューの潜在的便益の一つは、被評価者への建設的フィードバックとそれに続く被評価者の研究行動の改善である。被評価者のピア・レビューに対する認識および研究行動を改善することにおける評価の価値を確かめるためには研究がなされるべきである。或る革新的研究 [Luukkonen, 1993] では、ピア・レビューが被評価者の視点から扱われている。しかし、評価者から被評価者への情報移転の仕方を改善し、評価の勧告が改良された研究に転化する事を保証するためには、はるかに多くの研究がなされなければならない。



参考資料 4. 現在の評価関連人材に関する問題点の指摘例

明示的な評価システムの導入が遅れた我が国でも、様々な評価制度が設立・運用され始め、経過と共に評価の経験が蓄積し背景的にも理解は進んできたとはいえ、評価システムが信頼をもって機能するには、様々な課題がある。

なかでも、近年、評価制度が拡充されるに伴い、評価関連人材の量質の問題が急浮上している。これまで評価制度の導入に関心がややもすれば集中してきたといえるが、システムの信頼性に関わって人材問題が焦点の一つとなってきた。

評価人材の育成は、我が国の評価対象分野を問わず問題になってきている。日本評価学会第2回全国大会（2001年9月）でも「これまでの主要な関心は評価システムの設立と評価手法の探求に集中する傾向があった。しかし、評価が意味を持つためには、評価結果が信頼性を獲得することが不可欠の条件であり、それにはまず評価を行う者が専門的・倫理的要件を十分備えていなければならない。そこで、いかにしてそのような評価人材を育成するかが問題化してきたわけである」（共通課題セッション“評価人材の育成”）と指摘している。

本報告書は、こうした背景的な状況のもとに、研究開発評価の関連人材の育成の進め方を検討したものである。評価人材の問題にも、研究開発評価という専門分野には、固有の課題が認められる。以下は、本調査で行った専門家ヒアリング（下に専門家の方々のお名前を記した）ならびに既存の評価人材関連調査から得た評価関連人材に関する我が国の問題点である。評価人材の類型別に、問題点の指摘と部分的ではあるが改善策の提言を示す。評価人材の類型は、報告書の本文を参照されたい。もちろん評価の局面やフィールドなどは多様であり、ここにあげた問題点の指摘には、普遍的なものから一部の現象に留まるものまで拡がりがあると思われる。また、指摘者のもつ評価人材のあり方についての具体的な切りだし方にも相異があり、総合的な視点で編集したものではなく、一つの情報集積体である。評価人材問題を全体としての的確に把握するためには、相応の実態調査が必要不可欠である。

◆ 本調査で協力をいただいた専門家

石黒武彦	京都大学大学院理学研究科教授
大滝義博	(株) バイオフロンティア・パートナーズ代表取締役社長
桑原 裕	(株) ホエールズウェブ・ドット・コム取締役
佐野睦典	イノベーション・エンジン(株) 代表取締役
西村吉雄	東京大学大学院工学系研究科教授
丹羽富士雄	政策研究大学院大学教授
林 正男	お茶の水女子大学理学部生物学科助教授 (白楽ロックビル)
松重和美	京都大学大学院工学研究科教授 VBL施設長、国際融合創造センター長
松田修一	早稲田大学大学院(MBA)教授 早稲田大学アジア太平洋研究センター教授
村上路一	(株) 宇宙情報技術研究所副社長
渡辺 孝	(財) 理工学振興会 理事

◆ 評価マネジメント人材

(注)評価マネジメント人材の問題点の指摘には、マネジメント側の責任として評価システム自体の問題点と不可分な形で提起されている内容も多い。また、評価人材の能力以前に、評価の対象である研究開発システムやプログラム・政策自体が、適切に設計されず、また初めから評価を想定していないことが多かったことも、評価の混乱の背景にあり、留意が必要である。

▶ 評価システムの混乱に評価マネジメント人材問題がある（全般）

プロジェクト/プログラムの評価や研究機関の評価のマネジメント人材が、評価の目的、必要性や意義、ピアレビュー・システムの特徴を十分理解していないために、また、評価のマネジメントの専門的能力が不足しているために、評価が信頼をもって機能しない、あるいは、評価システムが改善され「標準」化されないという問題が起きている。

▶ 評価マネジメント組織/機関のシニア人材・コア人材に関する問題

我が国では評価の推進・マネジメント側は、科学技術の質的側面の専門性が乏しいために、ピアレビューの運営者（委員長、議長）になることは殆どなく、事務局としてピアレビュー体制を支える役割をもつことが一般的である。この場合でも、評価で最も重要な要素は、レビューを推進する機関としてシニア・マネジメントが高い品質のピアレビューを実現すべく関与することが必要である。

- 評価マネジメント組織/機関で責任を持った行動と意思決定が可能なシニア・マネジメント人材が評価システムを十分に理解せず、また関与が弱い。
- 評価推進組織に主体性と価値観をもって評価のマネジメントを進める人材が少ない。

▶ 評価のマネジメント自体の問題点

評価のマネジメント人材には、所属する機関やプログラムの評価目的をよく把握して評価システムの設計や運営ができる程度の能力が必要である。その能力で初めて評価の複雑な局面で、目的合理性や原則性・柔軟性・機動性をもった評価制度の運営や、評価結果を活かして次の展開に反映させることが可能となる。この点で次のような評価制度自体の問題が、人材問題と結びついて提起されてくる。

- 本来、評価はフィードバックし次により良くするために使うものであるが、不信と対立、あるいは評価以前に感情的抵抗が生じている場合がある。
- 評価における評価者間、評価者―被評価者間の「対話」の重要性が認識されていない。
- 評価では相互信頼と相互成長が不可欠であることが理解されず、評価関係者間での評価者側が被評価者側より上位にある感覚が蔓延したり、直輸入の監査型の評価がなされていることがある。
- 何のために評価をするのかが曖昧になり、「評価が自己目的化」している。スケジュール主義や数合わせ主義、評価の形式的義務化が進行している。
- 全体の説明がないままに評価者が呼ばれ、ラフな評価表で審査、短期のミーティングで修正して、それで終了、やりっ放しのものがある。
- 事務局があまりにも「事務屋」すぎる。時間で決まり、やってもやらなくても同じ印象。

- 膨大な労力をかけて誰も読まない評価報告書が量産されている。
- 評価する側、予算をつけた側に責任が絡み内部批判を怖れて率直な議論ができない。
- 評価に顧客の視点が入っていない。
- 評価には不確実性や間違い、主観性が伴うものであり、評価が信頼されるために、評価者に規律を要請し、様々な修正メカニズムが働く基盤である、「透明性」が不可欠であることが理解されていない。
- 過度に研究開発人材に依存したり、逆に研究開発の実態や論理が軽視されることがある。
- 評価者を選抜・登用し、動機づけ、評価能力を発揮させるマネジメントが不全である。とくに、ピアレビュー・チームである「評価委員会」の設置方法や運営に問題がある。

▶ 評価者の人選・構成の問題点

評価の質は評価者の能力を上回ることにはできない。現実に評価を行うピアレビューアの選定が決定的に重要である。

- 評価者の選考が恣意的にならないための手続きや基準、選考責任が不透明である。
- 評価の死命を制する評価者の選定が推進部署の情報に依存している。これを補完する人的・情動的なネットワークの規模・密度が極めて薄い。
- 科学技術の質的側面を扱う評価者が担当分野の専門能力にとどまり（ないし特定専門家に狭まり）、適正な課業か判断する広域的な評価能力に乏しい評価者編成がある（専門的な応援団/対抗者になることがある）。
- 複数のプロジェクト等の比較に利用するピアレビューのパネル間、分野間の正規化・標準化が意識されて構成されていない（相互比較に人的なバイアスがかかる）。
- 評価者の年齢、性別、セクター、出身学校などにバランスを欠いているものがありはしないか。

▶ 評価者の活用上の問題点

- 評価者が準拠する評価基準が明瞭でなかったり、十分に個々の評価の目的・対象等に適合的でないために、評価者が混乱することがある
 - ・ とくに評価者に共通の評価にむかう目的や物差しを与えることが不十分である。
 - ・ 評価者による評価情報の専門レベルや確信レベルが考慮されない。
- 評価者の人数が多すぎる。議論の時間が短い。真剣でない評価者の多数決は問題が多い。座長も進め方や良い司会ぶりを心得ていすぎる。
- 評価者の過剰な負担の回避、評価者のインセンティブ、評価の費用対コストなどに配慮した運営がなされていない。
 - ・ とくに参加者の費消時間コストが過小評価されている。
- 評価者の倫理基準が高いレベルで維持されなければならないが、制度面でも運営面でも必ずしも徹底されていない。
- 評価者の匿名性が十分に配慮されていない（研究の本質に関する正直で率直な見解が必要なら評価者は評価マネジメント以外には匿名である必要がある）。

▶ 評価マネジメント人材の能力開発が必要

- 近年は評価の専門家を徐々に確保する施策も見られるが、評価組織の中核的人材が

我が国の行政のローテーション人事の中で専門性を高められない。

- 評価マネジメント部局の職員に評価に関する適切な研修プログラムを整備すべきである。そのため評価に関する必要知識や必要経験、必要能力を明らかにして構造化し、各レベルに対応した教育研修体系を構築すべきである。
 - ・ 教育研修では評価の手法レベルのことなど関心が狭く哲学が曖昧である。
 - ・ ケース教育により総合的な状況認識・判断を研修することが必要である。
 - ・ 評価システムの生成発展史や成功・失敗例などを理解する機会が必要である。
 - ・ 評価のマネジメント人材相互の交流研修が必要である。
 - ・ 機関等のトップ人材にも（にこそ）理解を深めてもらうことが必要である。

▶ 専従的な評価マネジメント人材を確保すべき

評価マネジメントの一貫性を担保する人材政策が必要だが、異動が頻繁で、長期プロジェクトの経過を良く知る評価マネジメント人材が不足している。

- プロジェクト/プログラムの当初の状況や推移を考慮せず、評価時点での見方（とかく勢いがある）が支配的にならないように、継続して一貫して関与するメンバーが必要である。

▶ 専門人材の導入や専門人材による支援が必要

- 評価部署や評価部署をサポートする組織の専門性をサポートする評価専門家を専任ポストを拡充して採用確保する必要がある。また、支援・提携・外注する外部専門機関・専門家を適宜調達活用すべきである。
- 研究開発キャリアをもつ人材から評価マネジメント人材に転身するパスを整え、適切に処遇すべきである。

◇ 社会経済性分析エキスパート／社会経済性評価レビューア

▶ 社会経済性の分析にあたる専門家が不足し、評価の専門性が弱い。

きちんとしたインパクト・モデルと構造的データに基づき妥当なレベルで社会経済性な分析を行うこと自体が不足している。能力開発には経験やケース教育が不可欠な分野である。先行国で蓄積され始めた社会経済性分析にかかるマネジメントやメトリックスが未整備なままである。

- シンクタンクや行政内部で蓄積されるべき分析エキスパートが系統的に育成・蓄積・活用されていない。
- 社会経済的な側面の評価の専門性、評価者の役割が明確にされていない。
- 科学技術の内的論理に沿った評価と社会的経済的な評価が明瞭に区分されず、科学技術専門家が社会側を付度して本来の専門家に代行して扱っていることが多い。
- 評価が困難な対象であるために、好き勝手、思いつき、思い入れをいう関係研究者が登用されている。
- 経済専門家が研究開発の実態やダイナミズムを知らずに単純な経済性評価手法を進める。
- 社会経済的な評価が未成熟なまま、評価が形式的政治的に進められている。
- 海外の専門家の招聘にはコストと文化・言語の壁があり、注意すべきである。
- 中長期的に研究開発キャリアからの転身させて社会経済性分析の専門家にする教育

プログラムを拡充する必要がある。

- ▶ (分析データを前提としない比較的単純な枠組みの下で)直観的な判断能力を活用して事業化や社会経済的側面を評価する現場では、適切な専門家が評価現場で活用されていない。
 - 直観的判断を支える現場を知る人材が評価に入っていない(少ない)。
 - 経験を持つ科学技術人材を登用したセニョリティーパネルでも対応できていない。
 - 評価できない「大物」が権威づけだけのために登用される。
 - 著名な研究者・技術者によるエminentパネルも十分評価できていない。
 - 役所から独立的でない評価者の登用は問題がある。
- ▶ (とくに)事業としての評価で適切な人材が登用されていない。
 - 企業人材でも研究所在籍経験だけの人材等は、必ずしも実用化可能性や社会側の評価の論理を知らず、シーズプッシュ型の評価になりがちである。
 - 事業化では研究の実用化プロセスと市場からの技術評価の双方の評価能力が必要であり、企業の中でも分かっている人が少ない。「実際に」新規事業サイクルに関与した人材を捜すべきである(雇用環境の激変で有能な人材も確保しやすくなるだろう面はある)。
 - アイデア勝負のソフト分野などを別にすれば経験(40歳以上)が必要ではないか。
 - 評価者が責任を持ち現場をフォローするやり方も(評価人材育成面からも)有効である。
 - ベンチャーキャピタリストがもつ包括的な経験や情報収集方法が有効である。ベンチャーキャピタリストやベンチャー経営者でリタイヤした人材や現役でも優れた評価者ネットワークや研究開発情報をインセンティブに活用して登用すべきである。
 - 様々な技術現場を訪問し比較してきた技術ジャーナリズムで育った人材も評価者として有効である。

◇ 評価者(科学技術ピア)

- ▶ 適切な評価人材を確保できていない。
 - とくに未確立の先端的な分野では研究者層が薄い。
 - 高齢者が多く萌芽期の新しいテーマの評価ができない。
 - 若い一線研究者を登用し小数の円熟した調整役をおく体制も検討すべき。
 - 本気でしっかり評価を行う評価者を登用する。
 - 公募で意欲と能力ある人材を登用するチャンネルを拡充すべきである。
 - 有償でも一定期間専心できる人材を確保する制度が必要。
 - サバティカル・リープ人材や識見・活力ある定年後人材の活用を考えるべきである。
 - 評価人材の評価実績をフォローし、適宜推薦や公募を含む人材拡充を行っておく人材データベース・システムを整備する必要がある。
- ▶ 評価が特定の人材に集中する。
 - 評価者が疲弊する。
 - 評価がおざなりにならざるを得ない(しかし、卓見は重要である)。
 - 専門が細分化した結果、利害関係者以外から人材を確保することが困難である。
 - 対象の評価ができるプロがいなくて有名な先生方ばかりを集めた評価体制がある。

▶ 適正な評価者が選ばれていない。

- 評価者の客観性や見識をチェックするシステムが必要ではないか。
 - 各学協会からのリコメンドによる選定が必要ではないか。
- 「評価者の市場」「評価者の評価システム」を整備すべきである（しかし処遇や支援システムが低水準で片手間な状態のままではまずい）。

【評価者の評価能力や評価における態度の問題例】

- 専門的深さはともかく、広い視野と高い見識が不足している場合がある。
 - けなすことが役割と勝手に思っている評価者もいる。
 - 大学人に特許や市場の動向などは分からない人が多い。
 - 大学人は基礎学問的に面白くないと徹底的に攻めがちである。
 - 大学人は本性的に自分の研究が一番であり他はけなす傾向がある。
 - 大学人にはピンポイントの発想で全体を見通せない人も多い。
 - 企業人では本人の実績を追跡し見極めることが必要である。チームの単なる一員という場合もあり、評価者にふさわしい見識が備わっていないことが多い。
 - 大企業にいる人はしばしば会社イコール自分となる。会社を背負う気持があると余計なことはしない。評価には必ず賛否がありビジネス上マイナスなことはしない傾向がある。
 - 企業人の場合には、役所に近い人、役所に強い人には注意が必要ではないか。それだけメリットを受けてきたことで制約の危惧があり、評価者としてはよく吟味することが必要である。
- ▶ (マネジメント) 評価者のマネジメントを的確に行い、評価者の評価能力を引き出すべきである。
- 我が国のピアレビューは真剣さに乏しいことがある。
 - ピアレビューが応援団化することがある。
 - ピアレビューで逆に競争意識から評価者の専門分野では過度に厳しい場合がある。
 - ピアレビューの運営が恣意的になったり、あるいは“政治”に曝される場合がある。
 - 発言が明瞭な説明なしに葬られることがある。書類は一貫して作成蓄積公開すべき。
 - 事前の意図的な関連情報の配布、声の大きな意見の支配など。
 - ボス支配、学閥など人的ネットワーク、過去の感情的人間関係が評価を歪める。
 - 我が国では社会的風土として評価に伴う嫉妬の感情に留意する必要がある。
- ▶ (マネジメント) 評価者に評価のインセンティブが不足している。
- 評価は学会員本来の責務であり名誉であったが近年は必ずしも参加しない人材がいる。
 - 評価者の評価活動がどのように活用されどのような結果になったか、情報が開示されることがなによりも前提的に必要である。
 - 形だけの評価に動員される、評価結果が活用されないのではインセンティブがない。
 - 評価活動・実績自体が社会的に適正に認知されていない、位置付けられていない。

- ・名譽職としての評価者もいるが、評価の専門家の地位・社会的ステータスが低い。
 - ・評価実績を所属学会等でも業績として重視することが一つの方向である。
 - ・専門的職業能力として評価を位置付ける、キャリアパスを支える認定制度も検討が必要。
 - ・レビュー・サイエンスが軽視されている。
 - ・評価者としての活動が自分の研究のメリットとなることが必要。例えば、評価自体が研究開発活動に有益であるような真摯な研究討議・情報交換の舞台となることである。
 - ・報酬も検討材料ではないか。
 - ・評価者に特権や優遇を与えることは問題、評価者と被評価者の格差は危険である。
- ▶ (マネジメント) 評価者には権限と責任を明らかにして専門評価に専心させるべき。
- ・評価者には専門以外の評価をさせない
 - ・評価者の確信のある評価情報とそうでない評価情報の区分をすべきである
- ▶ (マネジメント) 評価者のレベルアップのサポートが必要。
- 今は評価者も勉強しないと一線の問題は分からないことも多い。常時、国際的なネットワークを動かして最新動向を交換するような評価コア人材やサポート人材が必要である。
- ・評価者相互の闊達な議論や相互の評価情報の交換の場が不足している（議論によるバイアスに留意することも必要である）
 - ・最新の研究情報・技術情報・特許情報などのサポートシステムが必要である（評価者も努力が必要。提案者の説明文脈からしかアプローチできない状況は危険である）。
 - ・同一研究・技術分野の評価者相互の意見交換やアップトゥデイティングの機会の提供が望ましい
- ▶ (マネジメント) 評価人材への評価マネジメント側からの情報提供・オリエンテーション・ガイドラインが不足している。
- ・科学技術ピアに改めて評価に関する教育研修等を義務づけることは困難である（研究開発活動自体で多忙かつ、既にステータスもある人材であり、評価自体は社会貢献的活動である）。適切でもない（科学技術の質に関する専門評価をこそ期待されており、その方向でのガイダンスは必要であるが、評価に対する知見は常識的なものでよい）。
 - ・評価の目的や意義の共通の理解が必要である
 - ・半年に1回でも国の方針の説明会や評価に関する事例を紹介する機会があればよい。
 - ・評価基準について過度な振れが出ないように研修が必要である（例：人材アセッサー）
- ▶ (マネジメント) 評価者の倫理、責任意識が不可欠である。
- ・利害関係を定義し関係者を排除する原則を徹底すべきである。
（評価者の層の薄い日本では形式的厳格性は困難ではないか）
 - ・評価者が“身内”の評価をしているケースがあり、評価の信頼性を低下させている。

- アイデア盗用や守秘義務違反はあってはならないし、噂だけでも評価の信頼性を毀す。
 - 明確な倫理基準とその遵守体制の整備が必要である。
- ▶ (マネジメント) 評価者や評価過程全体の公開が原則である。
 - 情報開示を通じて的確な洞察を行った評価者や明らかに誤った評価事例が分かるような評価者評価を機能させるべきである。
 - 評価者の良い意味での緊張の効果だけでも公開すべきである(負の影響も考慮すべき)。

◆ 評価人材問題の基盤的問題点

- ▶ 評価先行国に比べて被評価プロジェクトの規模に対応して評価への資源投入が低い。
 - 予算と人員を増やして評価に相応の資源を投入する必要がある。
 - 一定規模の資源投入を確保して評価専門家の受け皿を確立することが必要である。
- ▶ 評価のマネジメント実務を担当している人材への短期の研修プログラムがない。
- ▶ 評価のマネジメント専門家に対する実践的な教育プログラムがない。
- ▶ 研究開発評価や研究開発マネジメントの研究を行っている(大学院レベルの)研究機関や研究資金(プログラム)が少ない。
- ▶ 評価を支える高度な専門家が少ない(広く政策・マネジメントにわたる専門人材)。
 - 高等教育機関に評価人材の育成コースが設置されておらず、人材ストックはほとんどなく、他の分野から転向した小数の研究者や実務的専門家が散在しているだけである。
 - 研修コースや転換パスもほとんど未整備である。
- ▶ 固有の研究開発手法やアプローチを開発するなどしている支援機関(シンクタンクなど)が少ない。
- ▶ 研究開発評価や研究開発マネジメントについての教育が研究者技術者教育に十分に織り込まれていない。
- ▶ 評価を社会諸活動に織り込み評価風土を成熟させることや、評価、評価者が社会的に重要という認知と処遇が必要である。