

平成 29 年度文部科学省委託調査

「文部科学省の「研究開発評価研修プログラム
教材」の改定に係る調査・分析」
報告書

平成 30 年 3 月



本報告書は、文部科学省の研究開発評価推進調査委託事業による委託業務として、公益財団法人未来工学研究所が実施した平成 29 年度「文部科学省の「研究開発評価研修プログラム教材」の改定に係る調査・分析」の成果をとりまとめたものです。

したがって、本報告書の著作権は文部科学省に帰属しており、本報告書の全部又は一部の無断複製等の行為は、法律で認められたときを除き著作権の侵害にあたるので、これらの利用行為を行うときは文部科学省の承認手続きが必要です。

文部科学省委託調査
「文部科学省の「研究開発評価研修プログラム教材」の改定に係る調査・分析」
報告書

目次

1. 調査分析の概要	1
1.1 調査分析の目的.....	1
1.2 調査分析の内容・方法.....	1
1.2.1 有識者で構成される検討委員会の開催.....	1
1.2.2 研究開発評価研修プログラム教材改定案の作成と課題等に関する調査・分析.....	2
1.2.3 成果報告書の作成.....	3
1.3 調査分析の体制.....	3
2. 検討委員会の議事概要	4
2.1 検討委員会（第1回）議事概要.....	4
2.1.1 開催概要.....	4
2.1.2 議事の概要.....	4
2.2 検討委員会（第2回）議事概要.....	10
2.2.1 開催概要.....	10
2.2.2 議事の概要.....	10
2.3 ケースメソッド研究会（拡大検討委員会）の概要.....	15
2.3.1 開催概要.....	15
2.3.2 議事（意見交換会）の概要.....	16
3. 政策動向及び新しい研究開発評価の調査等	18
3.1 調査の概要.....	18
3.2 調査の結果.....	18

「研究開発評価研修プログラム」テキスト

※テキスト内では頁番号を独自に振り直している。

1. 調査分析の概要

1.1 調査分析の目的

文部科学省においては、研究開発評価に携わる人材を育成し、研究開発のより一層の推進を図るため、研究開発評価人材育成研修等を開催している。これまで研修等の教材として、平成 18 年度研究開発評価推進調査委託費により実施した「研究開発評価の質の向上のための調査・分析」の結果を活用して「研究開発評価研修プログラム教材」を開発し、活用してきた。

一方、平成 27 年に国立研究開発法人制度が開始され、同年、独立行政法人評価制度が改正された。また、平成 28 年 1 月には第 5 期科学技術基本計画が策定され、それを踏まえて「国の研究開発評価に関する大綱的指針」（平成 28 年 12 月）及び「文部科学省における研究及び開発に関する評価指針」（平成 29 年 4 月）が改定されるなど、我が国の研究開発評価を取り巻く環境はこの数年間で大きく変化しており、それに伴う教材の改定が急務となっている。

そこで、本委託業務では、新しい時代に合わせた研究開発評価の普及推進を図るため、研究開発評価に係る人材育成研修やシンポジウム等において使用するための「研究開発評価研修プログラム教材」の改定案を作成した。

1.2 調査分析の内容・方法

1.2.1 有識者で構成される検討委員会の開催

本業務に対する助言・協力を得るため、文部科学省と協議の上、本業務にふさわしい専門家 5 名の委員で構成された委員会を設置した。委員は次の通りである。

<検討委員（五十音順、敬称略）>

※○は座長

安藤二香 科学技術振興機構 社会技術研究開発センター アソシエイトフェロー

小湊卓夫 九州大学基幹教育院准教授

高橋真木子 金沢工業大学イノベーションマネジメント研究科教授

○ 林隆之 大学改革支援・学位授与機構研究開発部教授

丸山亮介 理化学研究所経営企画部評価推進課長

検討委員会は、評価の実務者等を交えてワークショップ形式で実施したケースメソッド研究会 1 回を含め計 3 回開催した。具体的には次のようなものである。

表 1-1 検討委員会の議事(ケースメソッド研究会を含む)

	日時	主な内容
第 1 回	2017 年 10 月 25 日(水) 15:00-17:00	1. 調査の全体像及び検討委員会の進め方について 2. 教材の構成・内容等について 3. 参考となる優良事例について 4. 今後の予定について
ケースメソッド研究会 (拡大検討委員会)	2018 年 1 月 25 日(木) 13:00-17:00	ワークショップ形式でケースメソッドを体感するとともに、ケースの書き方について演習を行った。また、研究開発評価に係る実務者の能力向上、特に「国の研究開発評価に関する大綱的指針」及び「文部科学省における研究及び開発に関する評価指針」において焦点の 1 つとなっているプログラム評価の設計・実施能力を養成するための手法としてケースメソッドを用いることの可能性や課題について検討し、教材改定のためのフィードバックを得た。
第 2 回	2018 年 2 月 20 日(火) 10:00-12:00	1. 調査の進捗について 2. テキスト案について 3. 今後の予定について

1.2.2 研究開発評価研修プログラム教材改定案の作成と課題等に関する調査・分析

(1) 政策動向及び新しい研究開発評価の調査等

これまでの文部科学省における大学・国立研究開発法人における研究開発評価についての調査・分析や当研究所が過去に実施してきた調査等の結果をもとに、評価現場のニーズや優良事例を取りまとめるとともに、平成 18 年以降の研究開発評価を取り巻く環境や政策動向の変化、科学技術基本計画や大綱的指針、評価指針の改定内容等に着目して、挑戦的な研究開発の具体的な評価方法等、新しい研究開発評価に関する有用な情報について調査及び整理を行った。

(2) 「研究開発研修プログラム教材」の改定案の作成

(1) の調査・分析の結果やそれを踏まえた検討委員会・文部科学省の意見を十分に反映し、研修プログラム教材の改定案を作成した。教材案の作成にあたっては、平成 18 年度「研究開発評価の質の向上のための調査・分析」、平成 27 年度「研究開発評価に関わる人材の現状と育成に関する調査・分析」及び平成 28 年度「文部科学省における研究及び開発に関する評価指針」の活用状況と課題に関する調査・分析など、これまでの文部科学省研究開発評価推進調査委託事業等の調査結果を適宜活用した。

研修教材は、初級及び中級のプラクティショナーを主な読者層として想定し、基本テキストに加え、演習問題も開発した。演習問題については、「文部科学省における研究及び開発に関する評価指針」(平成 29 年 4 月 1 日文部科学大臣決定)においてとりあげられている 4 つの特筆課題を対象に、それぞれについて参考となりうる先進事例をもとに作成した。具体的には、「科学技術イノベーション創出、課題解決のためのシステムの推進」4 事例、「挑戦的な研究、学際・融合領域・領域間連携研究等の推進」2 事例、「次代を担う若手研究者の育成・支援の推進」2 事例、「評価の形式化・形骸化、評価負担増大に対する改善」2 事例の 10 ケースをとりあげた。

1.2.3 成果報告書の作成

1.2.1 の検討委員会の議事概要、1.2.2 の調査・分析等の結果及び「研究開発評価研修プログラム教材」の改定案を成果報告書として取りまとめた。

1.3 調査分析の体制

本調査研究は、以下の体制で実施した。

<メンバー>

平澤 洽	公益財団法人未来工学研究所	理事長／上席研究員
大竹裕之	同 政策調査分析センター	主任研究員
田原敬一郎	同 政策調査分析センター	主任研究員
野呂高樹	同 政策調査分析センター	主任研究員
山本智史	同 政策調査分析センター	研究員

2. 検討委員会の議事概要

2.1 検討委員会（第1回）議事概要

2.1.1 開催概要

第1回について、以下の要領で実施した。

日時: 2018年10月25日(水)15時～17時

会場: 未来工学研究所 会議室

出席者:

(委員) 安藤委員、小湊委員、高橋委員、林委員、丸山委員

(事務局) 田原、大竹、野呂、山本、平澤(未来工学研究所)

議事:

1. 調査の全体像及び検討委員会の進め方について
2. 教材の構成・内容等について
3. 参考となる優良事例について
4. 今後の予定について

2.1.2 議事の概要

(1) 調査の全体像及び検討委員会の進め方について

主な委員からの指摘事項や質疑応答の内容は以下の通りである。

- 受講者について、現在は大学職員がほとんどで、プログラム評価を学ぶべき府省、研究開発独法、資金配分機関からは来ない。初級は、本当に初級（的事項）が必要であることがわかった。現在行っている研修を10年続けていくべきかと問われると難しい部分がある。これを機に研修プログラムも再設計するような形で検討できるとよい。大学職員向け初級向け研修だと、内容が難しい部分がある。研修自体の構成や全体像を考えながら、教材を作っただけなのがよいと思う。
- 学習到達目標は言葉として理解できるが、現場が何を考えていて、何が必要とされているか、構造化して焦点を絞っていく必要がある。プロジェクトでもよいので、評価を行う際の手順を考えた場合、どこに情報があり、どのような分析が必要かといった手続きの課題や、ロジックモデル等を通じて、評価結果を組織としてどのように活用していくかという議論になる。このようなフェーズに分けて考える必要がある。初級者は評価システムの設計をできる立場にはないので、それに対応する必要がある。初級は理解が出来ている程度、中級は少し使える程度だろうか。

- 総務省の研修は、必要性、有効性、効率性から入る。文科省研修でも、研修後半でインパクト・ファクターの概念がわかったという感想もある。単語の定義が初級には必要。他方、研究開発評価推進検討会の議事録にあったように、初級は具体的な事例がないと理解が進まない。
- 研修プログラムを受講するのは法人評価担当か、研究支援担当の方である。大学の場合は、初級の初級が必要である。
- 大学は業務として法人評価があるので、それに向けて準備しているのが実情。また、大型プロジェクト等であれば、中間評価、終了評価がある。大学の評価に関わる職員のほとんども3~4年でローテーションするため、その意味で中・上級は難しいと考える。ただし、課長級は数は少ないが中級以上の内容にニーズはあると思う。
- 大学、国研、ファンディング機関、府省とあるが、レベルの設定にあたっては評価の枠組みを作る人、作られた枠組みに対応する人がいることを念頭に置く必要がある。大学の運営費交付金を配分できる人にとっては、科研費を獲得するための呼び水としての**Proof of concept**という指標があり目的目標がクリアであり評価で悩まない。それ以外の分野では、独自性を出せる事項があまり多くない。資金の使途についても、外部評価がそもそも厳格なので、自己評価の余地があまり残されていない。予算等の使途が厳格にフレーミングされているものについては、エクセルや基礎統計等の事務処理スキルの向上のほうが、研究管理という点では適切なのかもしれない。国研は、大学とあまり変わらないところもある。ただし、研究開発系のファンディングの外部評価の場合、設計が適切であるかを見るが、実態としてそこまで提案時に考えられているとは限らない（例：プランニング時のロジックモデル作成）。また、学際融合や人材育成の部分は未着手で、フレームを作ることが中・上級に関わらず必要である。ファンディング機関については、枠組みを設計する力にフォーカスして良いだろう。府省については、PDCAを回す力が担当レベルで必要であるが、中・上級では必要に応じた生データの抽出や、評価手法を使いこなす力などが挙げられる。
- 大学でも大きな課題である。中長期計画でどのようにステップアップしていくかを示しているが、道筋となる設計図はなく、学内構成員も受け取り方がそれぞれ異なる。その部分を最初の段階でどのように設計していくか。評価制度で動いている大学においても重要である。
- 国立大学の法人評価に関わっているが、法人評価に独自性を出すことがメインターゲットであるかは見えないところである。
- 今回の教材が網羅することを目指しているものか、機関評価は大学改革支援・学位授与機構でやっているものをベースとするか等、整理をした方がよいと思う。文科省の研修が全てを網羅すべきかも検討した方がよい。最低共有すべきものを伝える場とするなど、切り分けが必要に思う。
- 法人評価は、決まっている中でやること。特徴を出すことは考えにくいという話もある。
- 研修・教材は評価指針を浸透させる目的もある。

- 大学でも評価が効果的であったかどうかの検証がおざなりになっている部分がある。何のために評価を実施しているか、評価の手続きをしっかりとしていく視点があってもよいと思う。

(2) 教材の構成・内容等について

<話題提供>

- 研発法人で運用しているプログラムの評価は十分にできていない。被評価者は研発法人ではなく、政策的な枠組みを与えている文科省やAMEDなどではないかとも考えられる。
- 評価の方法の観点では、研究以外の活動やアウトカムなど定量化しにくいものをいかに把握するかが難しく、評価のあり方を模索中である。定量的な評価は、基本的にビブリオメトリックスで行うことが多いが、学会発表が中心でない分野もある。
- 評価に関するデータベース化、システム化はリソースの問題があり、なかなか着手できない。評価のシステムを自前で持って活用することはかなり難しいと思う。
- 評価専門部署もジョブ・ローテーションでやっているの、専門家を育成できていない。評価者に関しては研究分野の利益代表的な側面が出てくるので、批判的な専門家の選定が課題である。
- 評価の形骸化対策については、常日頃評価システムの見直しを行う必要性を感じているが、一度決めた枠組みは変えにくい部分がある。ケースバイケースで改善を図っていくと枠組みばかりが増えてしまうことではないかと思う。
- 独立行政法人では独法評価と大綱的指針に基づく研究開発評価を両方とも実施することとなっており、研究者側から見ると、“評価”と名前が付いたものを2回やらされているという負担感を感じている場合もある。
- 理研を例にあげると、評価の構造として、理研アドバイザー・カウンシル (RAC) があり、それに報告をする形で、研究センターごとにアドバイザー・カウンシル (AC) が設置されている。PIの評価は、ACでも実施している。研究者個人の評価は上長が行う。課題評価は、各センターのACが行っているが、クライテリアは本部が出している。主任研究員研究室でも、別の枠組みで課題評価を行っている。
- 理研の各センターは大学でいう部局レベルである。ACの評価に関する実務は、個々の研究センターを担当する事務職員が業務の一環として評価も担当しており、専門的な評価担当がいるわけではない。それらのスタッフは併せて、予算要求業務も行っており、評価業務を通じて得た各種情報が予算業務にも活用できる点で、評価が有効活用されているといえる。個々の研究領域の特性などを踏まえた評価という意味でも、理研のACは事例として優良であると思われる。理研全体を包括的に評価するRACは、より大きなスコープの評価を行っており、評価担当のスタッフが担っている。

<質疑応答>

- ACがプログラム評価に相当するものと考えられる。RACの評価は、独法評価にインプットされるものか。
 - ビブリオメトリックスのデータ等は共有できるが、評価の業務としては別に走っている。
- RACは外部環境の見通しを立てて戦略を立てるところ、ACはどのようにやっていくかの戦術を扱っているところという認識でよいか。
 - ACのミッションには研究計画の評価とともにセンターの状況の評価、例えばセンターとして若手をどのように育成しているか等が含まれてくる。
- 理研では実質的にはプログラム評価が行われていると考えられるのではないか。
 - 評価指針に記載されている「プログラム」と合わない部分もある。概念としてのプログラム＝プロジェクトと、実態として使われている言葉との間に乖離があると思う。
 - 何を持ってプログラムと呼ぶか。例えば、社会技術研究開発センターは一つの「研究開発領域」ないし「研究開発プログラム」を起点にプログラム評価を行っている。一方で、より上位の単位をプログラム評価の対象としてはどうかとの考えもある。プログラム概念が浸透していないことが各機関の評価の現場を対象とした研修、教材テキストを作成する上ではネックとなる部分かもしれない。
 - CRESTを単位としても、戦略を単位としても、すべてプログラムと呼ぶことができる。
 - 理研の各センターをプログラムとして捉えると、脳科学総合研究センターは文科省の脳科学委員会が上位としてあり、政策に紐づいていることを考えると、プログラム評価の被評価者は脳科学委員会ではないかといった議論になる。
 - 研修テキストを作る上では、切り分けが必要である。法人評価の話をしたが、教育評価では内部質保証という言葉で、外向けの評価に対応するよりは内部向けに活用する方向である。使えるところは、第三者評価でも活用する形態をとっている。法人の第三者評価では、アウトカム目標を求めているが、日本では英国のようなインパクト概念が理解しにくい。法人評価で求めているアウトカムを示して、どのようにすればデータを取得することができるか、といったことまで示すことが必要になっている。
- 大学は、部局単位で研究戦略を持たないため、理研のACのような構造はない。部局レベルでの評価は難しいところである。
 - 以前は大学でも部局単位で外部評価を求めていたため、評価を行っていた時期があった。海外の大学では、部局ごとにアドバイザー・ボードがあるので、日本でももう少しできるとよいと思うが、それを一律に求めることは難しい。
- 大学内におけるURAの位置づけを巡る議論の中で、資源獲得にポジショニングしないと組織に定着しないとの話が以前あった。次年度以降の概算要求のデータを持っている組織は、評価の中で重宝される。評価が資源獲得に直結する大きな意思決定や組織運営

の中で活用されるようになると、かなり進展すると思う。それらの事例を見せることは重要である。

(3) 優良事例の収集・整理等について

- これまでの議論をまとめると、教材改定の方向性としては、①組織として対応しなければいけない評価制度と絡めながら、そこで真に求められているものは何かをきちんと理解できるものであること、また、②そうしたあるべき評価対応を組織内部で自発的に行うことが重要であるという認識をもってもらえるようにすること、その上で、③具体的にどのようにやればよいのかを学べるものであること、というイメージでよいか。
- 優良事例を示すことが重要との意見が多いのではないかと。
 - 優良事例の収集・整理にあたって、委員各位には事例の紹介やケース化する際の相談をさせていただければと考えている。国内の事例で適切な事例があれば、紹介いただけると助かる。海外の研究開発機関や資金配分機関の評価事例は、事務局で蓄積しているので、大学などで取り上げるべき事例をご教示いただきたい。
 - 東北大学の部局の組織評価や、電通大の URA による研究力指標作成などはよいのではないかと。
 - 広島大学が学内評価指標を作成している。
- 「優良事例」は誰にとっての優良であるか。資金の出し手側からみた優良なのか、それとも自分たちのアクティビティを可視化して次のアクションにつなげているような事例をそう呼ぶのか。
 - 優良の考え方は後者になると思う。独自に実施しているプログラム評価の例はある。昨年度、総務省の委託で SCOPE 若手枠のプログラム評価を試行した。これは、法律に基づくものではなく、改善のための独自評価であった。このような事例が国内でもいくつか出てきているのではないかと考えている。
 - 大綱的指針の話ではないが、いかに質の高い自己評価を実施し、組織学習・改善に活かすかが重要である。優良事例を集める際に、評価の質を高めようと努力したり、評価結果を事業等の改善に活かしているという視点は考えられる。RISTEX のプログラム評価は試行錯誤で実施している。今年度も継続して改善を図っている。特定のプログラムだけでなく、センターに反映させることも意識している。やり方そのものが良いという視点よりは、評価の意義を意識して実施しているものも事例としてあるとよい。海外事例は良くできているが、「それは海外だから」としっくりこないことがある。現場の人が参考となるような国内事例があるとよいのではないかと。
 - 個別の事例というより、評価改善に向けた組織的取り組みがストーリーとして示せるような例があるとよいと思う。
 - 初級の研修もそうであるが、評価に対する固定概念を打破する事例の置き方を意識して欲しい。

- 政策評価で悩ましい経験をした。研究開発計画ではアウトカム目標を設定することになっているが、全然できていない。大学と府省では大きく異なるが、研究開発の場合の特異性を明確にしたテキストを考えて欲しい。
- ロジックモデル研修では、必ず、“現場の担当者では判断できない”ということが出てくる。米国でロジックモデル研修を受けた者の話では、資源配分の権限を持った人がセットでないと、ロジックモデルは作れないと聞いたことがある。ワークショップで作り上げる場合は、現場の当事者のみでは限界があると思う。
 - ▶ 一口に「研修」と言っても正解がある問題に取り組む初級の研修と、自分たちの今後のアクションプランを見据えたファシリテーション等までを含む中級の研修の場合とは明らかに違ってくる。「中級」では研修の中にロジックモデルの作成まで含まれるのではないかと思う。ロジックモデルは、研究者側に抵抗があり、最初にこじれるとうまくいかないケースが多い。テキスト作りと、研修、ロジックモデルの3つの議論になっているので、分けた方がよい。制約条件がないままに、ロジックモデル研修を行うのは難しいと思う。
 - ▶ ロジックモデルの教育・研修で、ケースメソッドを用いることは理解できる。ただし、教授法やテキストを確立していく必要があるだろう。ファシリテーターの考え方に依存する場合もある。誰に対して、どのような形で実施するかを考える必要がある。中・上級には、ロジックモデルのファシリテーションができる能力が含まれるかもしれない。研修は、初級者が参加することを考えると、グループごとに、課題評価、機関評価と分けて実施できるとよいのではないか。
 - ▶ ロジックモデルは、プログラム評価にはぴったりくるが、機関評価だと仮想の大学を設定しても道筋や焦点が定まらない部分がある。1か月後にワークショップが開催されるので試してみたいところがあれば、検討できるので考えて欲しい。
 - ▶ ロジックモデルの言葉には、2種類あり、自己点検評価のツールとしてのロジックモデルと、戦略形成ツールとしてのロジックモデル（ロジカルフレームワーク）がある。
 - ▶ 多くの場合、事前評価が適切に行われていない中で、中間・事後の段階でプログラム評価を実施することになり、壁に当たっている状況である。やはり、最初の段階で仮説を立てて、評価をしていくということがないと発散してしまう。近年では、プログラムの設計段階で、ステークホルダー分析をしながら、ロジックモデル等で仮説構築し、それに基づいて公募、中間評価、事後評価を行うといったやり方が、多くの関与者が関わるものであるほど求められるようになってきている。ロジックモデルを利用する文脈を、狭義の評価の枠組み内にとどめるのか、プログラム設計を含めたより大きな意味での研究マネジメントまでを含めたものにするのかによって異なるのではないか。

2.2 検討委員会（第2回）議事概要

2.2.1 開催概要

第2回について、以下の要領で実施した。

日時: 2018年2月20日(火) 10:00~12:00

会場: 未来工学研究所 会議室

出席者:

(委員) 安藤委員、小湊委員、高橋委員、林委員、丸山委員

(事務局) 平澤、田原、野呂、大竹、山本(未来工学研究所)

議事:

1. 調査の進捗について
2. テキスト案について
3. 今後の予定について

2.2.2 議事の概要

(1) 全体について

- 初級、中級のレベルについて、経験年数等を設定、定義するか。
 - テキストにガイドラインをつけるなど対応したい。
- 中級の初級あたりを対象とするのか。
 - 未来工学研究所がこれまで行政官や独法等の職員を中心的な対象として行ってきた研修会では、自己流で評価を行ってきた人を初級、評価の概念を理解している人を中級と捉えてきた。そうしたプラクティショナーの中には現在、遙か先の方まで進んでいる人たちもいる。テキストは、そうした幅広い対象に使えるものにする必要があるのではないか。
 - 研修の実施経験からは、参加者の大半は、大学職員であり、初級にも至っていない。大綱的指針の話もあるが、制度的にどのような枠組みがあるかを理解するところと、それ以降の段階として、評価論という流れが適当だと思う。
 - 初級をやめるのではなく、中級等にも広げていくイメージである。大学内では、自己評価を実施し、評価結果を外部者に見てもらう機会がある。たいていの組織に改革へのマインドはあるので、よりよい評価の方法論についてきちんと書いていくべきではないか。
 - 大綱的指針は、機関長が考える部分があるので、大綱的指針をベースとしながらも、評価実務者に沿う形が望ましい。
- ケースメソッドについて、ケースメソッド教授法を今の研修会に入れることは難しいのではないか。ケースを扱うことには意義があるが、突然集めた人達に半日で教授法を伝

授できるかという議論の余地がある。

- ケースメソッドを用いた研修は講師を含めて学び合う機会になるよいのではないか。
- 初級中級の話が出ているが、IR 関連の研修ではスキルマップを作成する際に議論になった。評価においては、基本的な知識と必要最低限のスキルを身につけることが一つの側面としてある。ただし、全体の制度設計の中で何のためにやっているかの理解が必要。一方、それぞれの組織には独自の組織文化があり、組織文化に沿った評価が求められる。その理解も重要であるが、これはテキストで教えられる話ではなく、ケースとして、お互いどのような組織文化があり、共有するかが課題となる。
- 大綱的指針に沿って書きすぎると、ノウハウ集になってしまいかねないので悩ましい。また、意外と記述が難しいのが研究開発課題の評価。文科省評価指針では、競争的資金、重点的資金、基盤的資金と分けて記載しており、後二者については特にプログラム評価や機関評価との区分があいまいである。また、大綱的指針では、研究開発課題のうち、複数課題を束ねたものや大型のものをプロジェクトと呼称することもあると書いてあるなど用語法が統一されていない。こうした問題に関しては、コラムとして記載するなど工夫したい。
 - 大綱的指針はテキストをまとめる際にどこまでベースとなりうるか。実態として、プログラムとはいえないものをベースにしている部分を、プログラムに近いものになるよう全体の統一感を示すことが必要である。
 - プログラム・デザインや途上段階でのリデザインの方法についても書き込んでいく必要がある。
 - 評価のための事前の準備を何もしないまま、やみくもに評価作業を実施しているのが初級者である。どのような準備をしないといけないかがよく分かるといった内容が含まれているのが初級者向けテキストであるべき。評価の手順や準備といったものが、組織文化の中で根付いているところはよいが、そうではないところが多いし、かつて根付いていたノウハウが人事異動に伴い衰退しているところもある。何を準備しないといけないか意識させる必要がある。
 - 指摘の内容には評価の制度設計の部分も含まれており、この部分は現状では上級であるように思う。
 - 執行部を対象とするか、係員を対象とするかにより研修内容は異なる。必要な情報を学内で整理する体制等はあると思う。現状の係員（初級）を対象にしながら、全体像（執行部）もわかる形で議論した方がよい。
- 係員が実施している評価がどれだけ大事な意味を持つかを理解してもらうことが大切である。重要性がわかると、どのように評価書に書くべきかわかるようになると思う。
 - 評価の意義に関して、JST の PM 研修でプログラム評価をとりあげた際には、プログラム設計段階で作成するロジックモデルは仮説であり、それを検証するのが評価である、これは研究において仮説を立てそれを検証するというプロセスと同様である、と説明したが、納得して受け入れてもらった。プログラム設計において、必要性や位置づけをみるのは研究において先行研究をレビューするのと同じであるという話も腑に落ちたようで、研究とのアナロジーで説明するとよいかもかもしれない。
 - 評価は本来、研究計画を立てる作業と表裏一体である。研究者の自己評価から離れ、

評価に執行部が関わると、評価を受ける研究者側が評価の枠組みを見ないまま、全体的話が進むケースがある。

- いきなりロジックモデルというと戸惑うかもしれない。ただ、なぜ、必要かというのと、施策で実現したいこと（理念）を確認するためであり、施策を実現する仕組みとしてロジックモデルが出てくる。最初に実現したいことが内部で合意が採れていないと異なるものになってしまう。実現したいものが明確になっていると、それをプログラム化することは、それほど難しいものではない。プロジェクトとの違いから説明しようとする、神学論争となってしまう。

(2) テキスト案について

- 資料に基づき、全体構成について説明。
- 報告書の全体構成のうち文科省指針の特筆課題4点については中・上級者を主な対象にアドバンスな事例としてとりあげる。「科学技術イノベーション創出、課題解決のためのシステムの推進」については RISTEX や鳥取大学の事例ほか海外事例を紹介したい。「挑戦的な研究、学際・融合領域・領域間連携研究等の推進」については全米科学財団や理化学研究所の AC による評価を、「次代を担う若手研究者の育成・支援の推進」については EU や総務省の競争的資金を例にしたいと考えている。「評価の形式化・形骸化、評価負担増大に対する改善」については、上記の事例でも関連事項がでてくるが、北陸先端科学技術大学院大学における自己点検・評価の検証やカナダ自然科学工学研究会による共同研究開発プログラムの評価をとりあげることを想定している。
 - ▶ 若手育成支援について、国内の大学でもいろいろと取り組まれているが、支援策としてのみ語られ、メンタリングも含めた研究評価の話だとなかなか意識されていない。読者に「評価」概念を狭く採られるような区分を設けて論じるのはやめた方がよい。
- 全体構成として、「はじめに」があった方がよい。いきなり政策の枠組みの話がでてくると読み手はモチベーションが続かないかもしれない。評価は間口が広い。プログラム・マネージャーが日常的に行っているモニタリングも、日々のマネジメント業務にフィードバックするものであるという意味で評価として捉えることができる。実務者にとって評価は意義があることを触れた上で、概念や事例へと展開した方がよい。現実には、経験年数だけが多い初学者も多いのではないかと。多くの評価を経験しているが故に、評価を狭い範囲で捉えてしまう人もいだろう。評価を行うことのある種の面白みを最初に伝えた方がよい。
- 読み手としては、大学執行部も対象になるかもしれない。大規模な資金が配分され不祥事が起こるケースがある。評価の観点で言えば、モニタリングが弱いからである。年々、モニタリングを実施する体制は重要になっていると思う。近年はモニタリングは組織のトップすら対象に行われている。例えば大学の学長選考会議では、選んだ学長を評価の対象にするところが多い。
 - ▶ 近年みられるような学術研究に人材育成、産学連携も付加したタイプの巨額の研究プロジェクトの場合、事務体制が整っていないと研究費の扱いで不祥事等の問題になりかねない。こうした案件に研究者は特定期間のみ研究特区的に関わることで多いので、現場の実務者が適正な研究評価と資金の運用がきちんとわかっていないと

問題になると思う。そのようなことを学べるとよい。

- 年間 10 億円以上の大型研究課題では、研究費の使途や研究計画に対するモニタリング会議等のよほどの体制が整っていないと不祥事発生の可能性は常に存在する。現実の問題になっている不祥事案では、モニタリングに関する最初の制度設計が弱かったことが原因であるケースが多い。理研は従来からあったレビューと新しく加えた独法評価の組み合わせが抜群によくそれが効果的なモニタリングにつながっている。
- 「モニタリング」に関する事例はあるか。概念ではなく、事例としてあるものを意識的に入れないと、事前、中間、事後といった時期だけやればよいと思ってしまうかもしれない。
 - 固有名詞を出さなければ、多くの評価事例はあるが、事例として示す場合に固有名詞が必要となると使える事例の数は限られる。
 - モニタリングの事例として、RISTEX において日常的に行っているマネジメントは対象とならないか。
 - 学会発表等で Web 公表している。そうした公開情報をベースに公表するのは問題にならないと思う。ステークホルダーと協働していることもあり、個別プロジェクトのモニタリングのような評価結果を外部と共有するのは難しい側面もある。
 - モニタリングの種類も色々あると思う。モニタリングの時にインプット→アウトプット→アウトカムのロジックモデルを念頭に置けば、組織内で行われている活動の中で、この部分のインプットが行われていないとこのアウトプットが出てこないといった流れを把握できる。米国にはそうした事例が豊富にある。
 - 理研 AC の好事例としての特徴は、実務者による継続的なモニタリングを実施している点が挙げられる。AC ではピアレビューによって研究者のアクティビティを評価すると同時に、評価担当者が継続的に行っている研究者の評価結果を活用している。こうした二方向の評価を行っている例が理研の評価にはあり、日本には他に例はないのではないかと。こういったやり方を今のテキスト教材に実例として盛り込むのは難しいかもしれないが、デフォルトして掲載する方法もあると思う。
- 評価は、あるフォーマットに落とし込んで終わりというものではなく、ゴール達成のためのプロセスであり、その中でアウトプット、アウトカムがある。それがうまくいく場合といかない場合の中で、評価が行われる。その点では、どのようにこの部分を盛り込むかが鍵となる。学内の資源配分の決定者にも参考となる内容であるといえるかもしれない。ただし、評価実務者によっては重い部分であり、組織の中でどのように立ち回るかが示されるとよい。
 - 理想を言えば、マネジメントの視点を多少なりとも取り入れた事例があるとよい。
- 大綱的指針に即した方がよいとしたのは、先ほどの議論（評価の意義）に触れられているからである。「はじめに」で書くのがよいかわからないが、大綱的指針に記載してある理念をどこかに記載する必要がある。現案はいきなり制度論から入っている状況である。
 - 大綱的指針を読むだけでは、読み手によっては、迷子になる可能性がある。目的等の軸で整理しなおすとよいかもしれない。その入り口としての「はじめに」が必要である。
- これまで、個々の研究者を把握することにより大学全体の状況がわかり、よりよい施策

を打つことで機関を変化させることができる、といった議論を行い、整理してきた。こうした視点を盛り込むことができれば、研修としてはやりやすくなると思う。

- 政策評価の説明をどの程度入れるべきか。これは対象者の置き方によるかもしれない。制度論はあまり面白くないので、これについては簡単に説明する程度が適当である。また、機関評価は、プログラム評価とは異なり、組織としての視点があるので、戦略形成等にも関わる。研究者の業績評価は研究評価としてのみなされ、本来関連していなければならない人事考課とは別個になされてしまっているのが現状である。ただ、現状が問題であるという認識自体は共有されているが、では具体的に何をすべきかということはアイデアがない。
 - 研究者の業績評価は、問題提起はできるかもしれないが、こうあるべきという規範を持っておらず、適切に記述することがなかなか難しい。
 - どのような文脈で教員の業績評価を論じるのか。今は例外的に人事評価として行われており、実態は研究者の自己評価を組織の長が追認しているにすぎない。どのようなロジックで自己評価を組織評価と関連付けるか判然としない。研究者の業績評価について詳しく記述するのは個人的には難しいと思う。
 - 大綱的指針で研究者の「業績」という表現で扱われているが、メンタリングの話が含まれるのかも定かではない。本来、教員や人事の評価にはいろいろな視点があり、これらをまとめて整理して書くと言うことになじまない面もある。
 - どのような組織を作るかの観点で評価のポイントが決まっているのではないか。日本の大学では、大型の資金を得るために研究グループを形成している。研究グループの評価は、内部ではできないので、外部委員に依頼しているケースが多い。
 - 理研では研究者の業績評価は最終的に経営執行部レベルに上げられるが、PIは上長によるレビューと、アドバイザー・カウンシル（ピアレビュー）での二種類の評価が行われる。経営面で重要なのは、研究者の業績を評価・把握することで次の研究課題につなげていくことであり、芽が出ない場合は、別の課題を研究者に示すことがある。
 - 理研の事例は、業績評価のレイヤーで行った評価の結果を、階層の異なる上位の評価において重複なくうまく利用しているケースであると言えるかもしれない。そういう観点からも、理研の評価事例をテキストで紹介するのがよいと思う。
 - いずれにせよ、大綱的指針で十分に書ききれない項目が人事評価の部分であるとも言え、テキスト中でその点にも触れるとよい。
- 研修会のアンケート結果をみると、研修を受けてはじめてインパクト・ファクターの意義や限界が分かりましたといった反応が多く、ニーズが高い。現実をみても、意味を理解せずインパクト・ファクターで一律に評価させている状況があることも分かってきた。単一の指標ではない、その辺りのリテラシーを向上させるような記述が必要かもしれない。
 - 自己評価書を取りまとめる際に、文章の大元は大学教員が執筆するので、例えば研究評価について中途半端な知識を持っている医学系の教員が一律にインパクト・ファクターを用いて記載すると、誤った評価案になる。評価の多面的な要素を示すことができない。評価担当者として教員に指摘し、そのような事態を是正できるぐらいの知識がほしい。
 - アンケートからは、社会経済的なインパクトを入れることの必要性が理解され始め

てきている。このような要素も入れて欲しい。

- 評価をやっていると、確立されているものがあると思っている人がいる。自分たちの評価の目的に対して、どのようなものを取りいれるとよいか。そのような話と事例があるとよいと思う。事例の中で目的に対して、どのようなデータを取得しているかに触れられるとよい。
 - 日本が評価制度を採り入れた初期の頃には、一つのやり方を全てに適用するといったことが横行した。その後、複線的な評価が普及した。コラム的な形で多様に補填されていることが重要である。
 - 定型的な指標がなく、それを外部に求めても無理なので、自ら思考することが必要というメッセージを発するべきである。研修アンケートでは、評価者にどのような指標であれば評価をすることができるかと問うているが、この質問への答えはほとんどが寄せられない。答えのない問題でもある。そうであるならば、各機関の評価担当者には、自らの機関の研究の価値をどう評価して外部に発信していくかについてもっと知恵を絞ってもらう必要がある。これは皆で考えるべき問題で、2日間の研修会でどこまで理解してもらえるかという課題はあるが、教科書としては、そうした本質的なところを記載するという方針でいくとよいのではないか。
 - 総務担当理事には、研修会の案内をしてもよいと思う。
 - 初級向けの2日間の研修でできることには限界がある。中級向けの研修を別途考える等、割り切る必要がある。2日間できることは、より初級的なものを実施し、テキスト案は参考資料として位置付けることもあり得る。次のステップを示すガイドラインができるとよい。事例として大学が中心であると、大学で働いたことがない人にはわからない。
- 現在実施している研修においても、言葉の使い方がまちまちだったりする。用語を統一するためにテキストを活用することもありえる。その意味で用語集は重要。アウトプットとアウトカムがわからないという意見も聞く。手間がかかり、時間的に難しいかもしれないが、用語を説明した上で、それが本文中のどこに記載されているのかのインデックスをあわせて示せるとよい。
 - 政策評価法と大綱的指針で、アウトプット、アウトカムの使い方が異なるので、あるべき使い方として書くことはできると思う。
 - 研修は、講師が変わってもできるとよいと思っている。それに資するものにテキストがなればよい。

2.3 ケースメソッド研究会（拡大検討委員会）の概要

2.3.1 開催概要

研究会について、以下の要領で実施した。

なお、ケースメソッドとは、もともとハーバード・ビジネス・スクールで開発された事例研究法・教育手法の1つであり、実際に起きた事例を教材に、現実の諸問題について集団で討議を行い、問題解決能力等の向上を図るものである。本研究会では、ハーバード・ビジネス・スクールで日本のケース作成を担当されていた山崎繭加氏を講師にお招きし、実際の

ケースを用いて学習効果を体感してもらうとともに、ケースの書き方について学ぶことを目的とする。その上で、研究開発評価に係る実務者の能力向上、特に「国の研究開発評価に関する大綱的指針」及び「文部科学省における研究及び開発に関する評価指針」において焦点の1つとなっているプログラム評価の設計・実施能力を養成するための手法として、ケースメソッドを用いることの可能性や課題について議論した。

日 時:2018年1月25日(木)13時~17時
会 場:金沢工業大学 虎ノ門キャンパス 13階(1301)
プログラム:
13:00-13:10 イン트로ダクション
13:10-14:15 「Cafe はまぐり堂」ケース議論
14:15-14:30 休憩
14:30-16:00 ケースの書き方について 講義・演習・Q&A
16:00-17:00 意見交換会

2.3.2 議事（意見交換会）の概要

(1) ケースメソッドの必要性について

- 日本においては、府省でも国立研究開発法人でも大学等でもプログラム評価が定着しておらず、手がかりとなるものはガイドラインとしての大綱的指針や文科省評価指針しかない。そのような状況下で、自己点検・評価を基本としたプログラム評価を推進していく必要がある。
- 決まった枠組みのない中、自己点検・評価を行うには、評価実施者自らがプログラム評価のためのシステムや枠組みを考え、構築する必要があるが、プログラム評価を行うことの意義もよく分からない中で進めるのは非常に困難。
- プログラム評価を実質化するには、(初級者にせよ中級者にせよ) 頭で考えて理解するだけではなく、その意義を含めて感覚的に理解する必要がある、それを実現するための最もよい学習方法の1つがケースメソッドではないか。
 - ケースメソッドは「(問題) 状況を整理し、問題は何なのか、もし自分が当事者として責任ある立場に立っていたら、どのようにその問題を解決するか」(百海 2009) をケースに基づき考える方法である。
 - 「マネジリアル・スキル」(Katz)のうち、「コンセプチュアル・スキル」を養成することに主眼。研修方法として、グループ討論等を取り入れることで「ヒューマン・スキル」の養成にも寄与。「テクニカル・スキル」についてはテキストで学ぶことができる。

(2) ケースメソッドの活用方法及び対象について

- ケースで学ぶ目的としては、次のようなことが考えられる。

- (プログラム) 評価の意義を理解してもらう。
- 事業及び(プログラム) 評価のシステムを考える上での重要な要素を理解してもらう。
- 組織の生き残りのために、どのようなことを対外的に示していかないといけないか？
- 組織のミッションを果たすために、どのように事業を作り込んでいけばよいか？
- 事業の改善のために、どのような評価の仕組みをもたないといけないか？、等
- 研究開発評価研修プログラム教材における位置づけとしては、そもそも「答え」のない問題を対象にする。具体的には、文科省指針における特筆課題に対応したものにする。これらは多くの主体が悩んでいる課題でもある。
 - 科学技術イノベーション創出、課題解決のためのシステムの推進
 - 挑戦的な研究、学際・融合領域・領域間連携研究等の推進
 - 次代を担う若手研究者の育成・支援の推進
 - 評価の形式化・形骸化、評価負担増大に対する改善
- 一方、ストーリー仕立てで詳細な文脈まで含めてケース化するのは、とりあげる機関の承認等が必要であり、短期間でまとめるのは難しい。海外事例を含め、公開情報をベースにするしかない。

3. 政策動向及び新しい研究開発評価の調査等

3.1 調査の概要

これまでの文部科学省における大学・国立研究開発法人における研究開発評価についての調査・分析や当研究所が過去に実施してきた調査、検討委員会等の結果をもとに、評価現場のニーズや優良事例を取りまとめるとともに、平成 18 年以降の研究開発評価を取り巻く環境や政策動向の変化、科学技術基本計画や大綱的指針、評価指針の改定内容等に着眼して、挑戦的な研究開発の具体的な評価方法等、新しい研究開発評価に関する有用な情報について調査及び整理を行った。

具体的には、「文部科学省における研究及び開発に関する評価指針」（平成 29 年 4 月 1 日文部科学大臣決定）において特筆課題として示されている 4 課題を考える上で参考となりうる国内外の事例を幅広く収集するとともに、ケースメソッドの素材として相応しいものを 10 事例抽出し、その結果をテキストの第 4 章でとりまとめた。

3.2 調査の結果

テキストにおいて、最終的にとりあげた事例及び着目点は次のようなものである。

表 3-1 課題別の事例と着目点

事例	着目点
特筆課題 1: 科学技術イノベーション創出、課題解決のためのシステムの推進(4 事例)	
オランダ科学研究機構応用・エンジニアリング科学領域によるオープン・テクノロジー・プログラムの評価	科学的な挑戦性と成果の社会的有用性をどう両立させるか？(プロジェクト評価)
鳥取大学による地域参加型研究プロジェクトの評価	地域課題の解決に貢献しうる研究をいかに促進できるか？(プロジェクト評価)
科学技術振興機構社会技術研究開発センターによる研究開発領域の中間評価	ステークホルダーの実質的な参加と研究者の意識変革を促すためにはどうしたらよいか？(プログラム評価)
Nesta(UK)によるビッグ・グリーン・チャレンジの仕組み	横展開のためのモデル化と政策提言としての評価(プログラム評価)
特筆課題 2: 挑戦的な研究、学際・融合領域・領域間連携研究等の推進(2 事例)	
全米科学財団による INSPIRE イニシアチブにおけるプロジェクト評価	ピアレビュー以外の評価方法の模索(プロジェクト評価)
理化学研究所におけるアドバイザー・カウンスルによる機関評価	外部評価を活用した世界最先端の研究開発を推進する機関の評価(機関評価)
特筆課題 3: 次代を担う若手研究者の育成・支援の推進(2 事例)	
欧州研究会議による若手研究者支援のための競争的資金 Starting Grants におけるプロジェクト評価	自立した若手研究者を育成するための評価基準とは？(プロジェクト評価)
総務省 SCOPE における若手 ICT 研究者等育成型研究開発に対する試行的なプログラム評価	特定分野の人材を育成・確保するためのプログラム・デザインとは？(プログラム評価)
特筆課題 4: 評価の形式化・形骸化、評価負担増大に対する改善(2 事例)	
北陸先端科学技術大学院大学における自己点検・評価の検証	自己評価能力を向上させるための仕組み(機関評価)
カナダ自然科学工学研究会議による共同研究開発プログラムの評価	調査・分析・評価と意思決定を有機的につなげる評価システムの構築(プログラム評価)