

ISSN 1346-6151

日本評価研究

Japanese Journal of
Evaluation Studies

Vol. 20, No. 1, March 2020

研究ノート

評価倫理における「評価協力者の尊重」の検討
—心理学分野の研究倫理からの示唆—

小林 信行

庁が実施する評価の現状と課題

鏡 圭佑

春季第16回全国大会開催報告

委員会活動報告

日本評価学会
Japan Evaluation Society

『日本評価研究』編集委員会
Editorial Board

編集委員長
Editor-in-chief

山谷 清志(同志社大学)
Kiyoshi YAMAYA

副委員長
Vice-Editor-in-chief

西野 桂子(関西学院大学)
Keiko NISHINO

常任編集委員
Standing Editors

牟田 博光(東京工業大学)
Hiromitsu MUTA

編集委員
Editors

岩渕 公二(NPO政策21)
Koji IWABUCHI

大島 巖(日本社会事業大学)
Iwao OSHIMA

岡本 義朗(鳥根県立大学)
Yoshiaki OKAMOTO

小野 達也(鳥取大学)
Tatsuya ONO

窪田 好男(京都府立大学)
Yoshio KUBOTA

佐々木 亮(国際開発センター)
Ryo SASAKI

田中 弥生(会計検査院)
Yayoi TANAKA

南島 和久(新潟大学)
Kazuhisa NAJIMA

西出 順郎(明治大学)
Junro NISHIDE

林 薫(文教大学)
Kaoru HAYASHI

源 由理子(明治大学)
Yuriko MINAMOTO

〒108-0075 東京都港区港南1-6-41 品川クリスタルスクエア 12階
一般財団法人国際開発センター内
特定非営利活動法人日本評価学会編集委員会
E-mail: journal@evaluationjp.org

日本評価研究

第20巻 第1号 2020年3月

目次

研究ノート

小林 信行

評価倫理における「評価協力者の尊重」の検討

—心理学分野の研究倫理からの示唆—..... 1

鏡 圭佑

庁が実施する評価の現状と課題..... 15

春季第16回全国大会開催報告

開催の報告とお礼..... 27

プログラム詳細..... 28

共通論題セッション報告..... 30

自由論題セッション報告..... 34

Brown Bag Lunch Session 報告..... 37

委員会活動報告

企画委員会..... 39

研修委員会..... 39

出版・広報委員会..... 40

国際交流委員会..... 40

日本評価研究刊行規定..... 42

日本評価研究投稿規定..... 44

日本評価研究執筆要領..... 46

日本評価研究査読要領..... 49

Publication Policy of the Japanese Journal of Evaluation Studies 51

Information for Contributors (For English Papers) 53

Writing Manual of the Japanese Journal of Evaluation Studies (For English Papers) 55

Referee-Reading Guideline..... 57

【研究ノート】

評価倫理における「評価協力者の尊重」の検討 ー心理学分野の研究倫理からの示唆ー

小林 信行

OPMAC株式会社¹
kobayashi@opmac.co.jp

要 約

本稿は、心理学分野の研究倫理における研究協力者への配慮を視座に、評価倫理における基本原則「人々への敬意」に関する論考を行う。そして、その論考を踏まえ、日本評価学会の「評価倫理ガイドライン」で今後反映すべき事項に関する示唆を導出する。心理学分野では、研究は実践活動よりも高リスクと見なされ、研究に先立ち倫理審査が求められる。同様に、評価協力者に著しいリスクが生じる可能性がある場合、類似する審査が望まれ、その要否の判断基準が検討課題となる。「評価倫理ガイドライン」は、評価が介入の割り付けを決定する状況を想定しないため、前記の状況でのインフォームド・コンセント、許容される統制群の設定も検討課題となる。研究倫理の「実践」から、幾つかの課題も明らかとなった。倫理審査が調査手法の選定に影響し、特定の手法を忌避する傾向が見られた。また、「理論」が提示する複数の基本原則どうしが衝突し、倫理的な判断が困難となった結果、倫理面の配慮が研究協力者の負担を重くするケースも生じている。

キーワード

評価倫理、研究倫理、倫理審査、インフォームド・コンセント、統制群

1. はじめに

社会科学の手法を使う評価者は顧客を持つ専門職（consulting professionals）と研究者（scholarly professionals）の中間にあるとされる（Fitzpatrick 1999, p.8）。後者である研究者の倫理においては、その内容は研究公正（research integrity）、研究倫理（research ethics）に区分される（Steneck 2006, p.56）。前者の研究公正は知識の生産者としてのマナーであり、ねつ造・改ざん・盗用といった不正行為を扱う。後者の研究倫理は研究活動と社会の倫理的価値との合致を問い、研究協力者

の保護、動物実験、研究成果の軍事利用等の倫理面の問題を扱う。アカデミアの世界では、医学研究を筆頭に研究協力者を搾取する危険性があり、その保護は研究倫理の主題となってきた。

日本評価学会の「評価倫理ガイドライン」における基本原則「人々への敬意」は評価に携わる者が重視すべき価値であり、評価実施に際して情報提供者、受益者を含む関係者の安全・尊厳への配慮が求められる。つまり、評価倫理における「人々への敬意」は前述した研究協力者の保護と通底する論点を有している。しかし、評価と研究が実施される環境には相違点もある。

特に医学研究の研究倫理では、リスクの高い身体への直接的な侵襲が主に扱われ、相対的には低リスクで、かつその幅が広い評価分野との社会介入とは扱うリスクの性格がやや異なる。

そこで、本稿は評価倫理における「評価協力者の尊重」への貢献を目的とし、研究倫理における配慮事項との比較を行い、「評価倫理ガイドライン」に反映すべき事項に関する示唆を導出する。なお、本稿は比較対象とする研究分野として心理学を選定し、日本心理学会、日本教育心理学会、日本心理臨床学会の倫理規程を参考に検討を進める。

2. 本稿の視点と分析方法

(1) 心理学分野の研究倫理と比較する理由

日本評価学会の「評価倫理ガイドライン」の課題として、「人々への敬意」が基本原則に含まれるものの、評価協力者の保護に関する行動指針が限定的な内容であることを挙げたい。まず「評価倫理ガイドライン」は、情報収集における個人情報保護、自由な発言を担保する配慮、インフォームド・コンセント等を扱うが、評価が介入を割り付ける状況は扱われていない（日本評価学会 2012、p.6）。近年、国内の評価分野において、ランダム化比較試験を導入する動きがあり、評価自体が介入の割り付け方法を決定し、評価協力者に過少な便益、時には負の便益をもたらすリスクが高まっている。

次いで、「評価倫理ガイドライン」は評価協力者が直面するリスクの高低を想定せず、一律に配慮を求めるのみである。実際には、情報収集の手段や収集する情報の内容次第で、評価協力者のリスクの程度は異なるはずであり、評価毎にリスクとその対応を精査する必要がある。

加えて、「評価倫理ガイドライン」は評価協力者が多様である点を考慮していない。「評価倫理ガイドライン」のインフォームド・コンセントでは評価協力者が評価内容を理解し同意できることを前提とする。しかし、評価対象となる介入次第で（例えば、教育や社会福祉プログラム等）、この前提は不適切となる状況も生じる。

上記した課題には、昨年改訂された米国評価学会（American Evaluation Association: AEA）の倫理規程 Guiding Principles for Evaluators（以下、Guiding Principles）も答えを与えない（AEA 2018）。そこで本稿は学術研究に着目し、様々な状況における配慮事項を抽出し、評価分野への適否、導入時の留意点を論考する。

学術研究のうち、本稿では心理学を選定する。その理由として、心理学研究にはネガティブ・リスクのある心理療法の効果測定も含まれ、その介入の割り付けに長年関与してきた。また、心理学研究はリスクの高い行為（心的外傷に触れる質問、だまし行為等）を伴うこともある。心理学は研究協力者のリスクへの対処という点では他の社会科学分野に先んじている。同分野の研究協力者には、自律的な意思決定が難しい層も含まれ、この点からの示唆も期待される。

心理学と比較し論考するメリットとして、同分野のプロフェッションが研究者と実務家から構成される点も指摘したい。特に本稿で比較対象とする3つの学会には、研究者のみならず、心理学を実践に用いるプロフェッション（心理療法家、スクールカウンセラー等）も会員に含まれる。

つまり心理学分野の倫理は異なる2つの系譜を有する点が特徴的であり、評価の源流となる他の学問領域（経済学、社会学等）とは異なる。特に各学会の倫理規定では研究以外の実践活動の倫理も扱うため、研究と実践活動双方の倫理を比較し、研究倫理の相対的な立ち位置を提示できる。また、研究者と実務家は所属する組織の性格の違いから、利用できるリソースが異なり、プロフェッション内のこの差異への対処を考察することも可能となる。

(2) 用語の定義

心理学分野では、研究対象となる者に、主従関係を示唆する恐れのある「被験者」以外の用語を使う。具体的には、日本心理学会の「倫理規程」では「研究対象者」を使い、研究分野別に用語を言い換える場合、調査研究では「調査協力者」、実験研究では「実験参加者」、フィールド研究では「研究協力者」の用語を充てる。

そこで本稿は統一的な表記として、同分野の研究倫理の用語のうち、「研究協力者」を用いる。

「評価倫理ガイドライン」では「調査対象者」の用語を使うが、その範囲は介入の対象者以外も含む。そのため、本稿では論点を明確にするべく、作業定義として評価が実施される介入の対象者に限定して「評価協力者」の用語を充てる。

各学会は倫理上の基本原則や行動指針を公式に文章にまとめたものに対し、倫理綱領、倫理基準といった様々な用語を充てている。本稿では、職業上の自律的規範を提示する文書に一括して「倫理規程」という用語を充てる。また、個々の倫理規程に言及する場合、固有の呼称に「」を付けて区別する。

(3) 分析の枠組みと手順

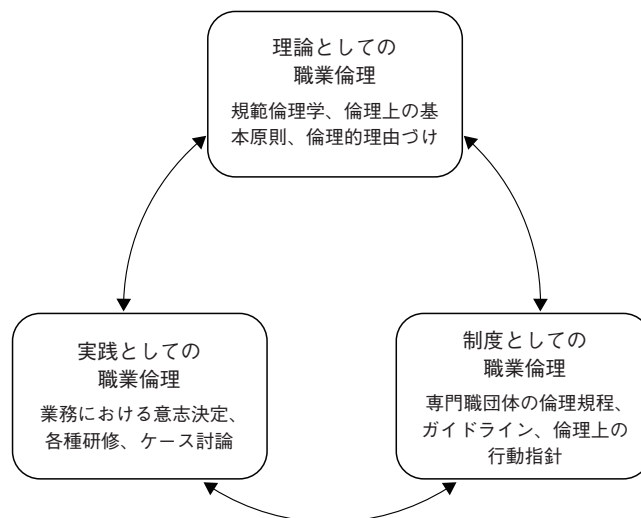
梅津（2002）は、職業倫理を包括的に捉える概念的枠組みとして、職業倫理を3つの側面から説明するフレームワークを提唱している（p.7、pp.128-130）。このフレームワークでは、日常の倫理の意思決定と行動である「実践」、それを省察し意思決定の理由づけの根拠となる「理論」が存在する。「理論」には倫理的判断において重視されるべき価値を示す基本原則が含まれる。

そして、「理論」が提示する基本原則を「実践」における個々のケースに整合的に適用する機能として「制度」が「実践」より外化してきたと説明している。この3つの側面が相互にフィードバックを繰り返し、職業倫理は発展し続ける。

本稿冒頭で言及した通り、評価者は社会科学の調査手法を使うという特性を有する。そのため、本稿は研究倫理における研究協力者への配慮、研究倫理の「実践」からの課題から評価分野への示唆を導出するが、あくまでも評価プロフェッションの職業倫理が主題である。特に研究では想定されない評価分野の「制度」上の特徴が研究倫理の直接的な適用を妨げる可能性があるため、研究倫理の評価分野への適否を検討するに際し、両分野での「制度」の違いを考察する必要がある。研究倫理のフレームワークでは「制度」を明示的に扱わないため、本稿は概念的枠組みの中に「制度」を明確に位置づける前述の梅津の職業倫理のフレームワークを選定した。

八田・町田（2003）は梅津のフレームワークをさらにプロフェッションの職業倫理へと展開し、会計士の職業倫理のモデルを策定した。そのモデルでは、日常的な職業倫理の「実践」とそれを促進する倫理教育（会計事務所内の研修、

図1 プロフェッションの職業倫理フレームワーク



（出所）梅津（2002）p.7、八田・町田（2003）pp.42-43に基づき筆者作成

専門職団体の継続教育等) (pp.42-43) が「実践としての職業倫理」に位置づけられる。次に「理論としての職業倫理」では、規範倫理学が職業上の価値を考察し、倫理原則、倫理的理由づけの根拠を提示する。最後に「制度としての職業倫理」において、職業倫理を個別事例に整合的に適用するために、明文化された行動指針(専門職団体や会計事務所内の倫理規則、関連法規等)が策定される。これまでの議論から、プロフェッションの職業倫理には、図1に示されるフレームワークが想定される。

このフレームワークに沿って、本稿はまず倫理規程を中心に「制度」を論考し、次いで「実践」から「理論」や「制度」へのフィードバックを扱う。具体的には、本稿で扱う心理学諸学会の倫理規程は研究と実践活動の双方を扱うため、最初に両者を比較し、研究倫理の相対的な立ち位置を示す。次いで、心理学分野の研究協力者の尊重に関する事項につき、評価分野への適否を論考する。最後に研究倫理の「実践」から見えてきた課題を提示し、評価倫理の「理論」や「制度」への示唆を導出する。

前述の通り、「理論」には倫理的判断の根拠となる価値を示す基本原則が含まれる。AEA(2018)のGuiding Principlesの前文では、評価において複数の基本原則が対立しうる点に着目し、またWeiss(1997)は「人々への敬意」が他の基本原則と抵触する点を示唆する(p.430)。そのため、本稿では「実践」から「制度」へのフィードバックにて、研究倫理の現状を踏まえ、評価分野で「人々への敬意」を巡って発生が想定される課題を抽出し、その対応を検討する。

3. 今日の研究倫理に至る経緯

心理学分野の研究倫理から示唆を導出する前に、本稿の議論を理解する一助として、心理学分野の研究倫理の経緯を説明したい。

研究倫理における倫理規程は1947年のニュルンベルク綱領を原点とするが、米国の生物医学・行動研究における被験者保護のための国家委員会が1979年に発表したベルモント・レポートが

表1 ベルモント・レポートの構成

前文：人を対象とする研究の倫理原則とガイドラ

イン

パート A. 診療と研究の境界

パート B. 基本的な倫理原則

1. 人格の尊重

2. 善行

3. 正義

パート C. 適用

1. インフォームド・コンセント

2. リスク・ベネフィット評価

3. 被験者の選択

(出所) 津谷ら(2001)の邦訳を参考に著者作成

今日の研究倫理に至る重要な分岐点と考えられている²。第1の理由として、ベルモント・レポートは現在の研究倫理の根幹となる概念を提示している。

表1に示す通り、ベルモント・レポートは倫理上の基本原則－人格の尊重(respect for persons)、善行(beneficence)、正義(justice)－を示し、研究にて尊重されるべき価値を明確にした(川瀬2013, p.62)。加えて、研究の範囲にも進展が見られた。医学分野で研究倫理の適用範囲を設定する上で、研究と治療の区分は根源的な課題である。ベルモント・レポート以前の医学研究では、研究者の意図を区分の基準とし、研究者自身が治療と研究の区分を行っていた(田代2011, p.102)。それに対し、ベルモント・レポートでは、治療を成功への期待を伴うものと位置づけて、研究者自身の判断ではなく、他の専門家や規制当局による効果の認識に基づき、研究と治療を区分する。ピアグループによる効果の容認を区分の基準とする形式は「承認モデル」と呼ばれる(田代2011, pp.95-96)。

第2の理由として、ベルモント・レポートが示したガイドラインは、今日における研究の倫理審査の基礎となっている(Oaks 2002, p.448)。ベルモント・レポートは、前文で既存の倫理規程が「複雑な状況をカバーするのに適切でないことも多く、ときに相い矛盾し、解釈や適用が難しい点も多い」と指摘し、その解決を企図した。

ベルモント・レポートは、インフォームド・コンセントや研究の便益とリスクにつき詳細な記述があり、倫理上の判断に対しガイドラインを提供する。

第3の理由として、社会科学の研究倫理に与えた影響がある。ベルモント・レポートは医学のみでなく、行動科学（behavioral science）も対象とした。ベルモント・レポートが提示した概念は米国連邦政府の省庁が実施ないし補助する、人を対象とする研究に適用される「コモン・ルール」（Title 45, Part 46, of the Code of Federal Regulations, 通称45CFR46）にも反映された。そして「コモン・ルール」は1991年以降、米国連邦政府より補助金を受ける研究機関や大学に適用され、政府補助金を受ける研究機関や大学で実施される社会科学研究もその遵守が求められた（Oaks 2002, p.448）。

上記した医学及び行動科学全般の流れに沿い、米国の心理学分野でも1970年代から研究倫理が発展してきた。具体的には、まずアメリカ心理学会（American Psychological Association：APA）が1953年に採用した初の倫理規定「サイコロジストのための倫理規準」（Ethical Standards for Psychologists）は新しいプロフェッションである心理療法家の倫理の手引きを目指すものであった（Sabourin 1999, pp.54-58）。APAにおける研究活動の倫理は、後年となる1972年の「人間の参加者を使用した研究行為における倫理綱領」（Ethical Principles in the Conduct of Research with Human Participants）の採用を待たねばならない。前述のベルモント・レポートの発表後、心理学分野でも研究倫理が強く意識され、「コモン・ルール」の幅広い研究分野への適用は研究倫理の順守を必須とした。

一方、日本における研究倫理は歴史的な文脈を欠いたまま、医薬品の治験の必要性から導入されたとの意見がある（笹栗・池松 2011, pp.9-11）。同様に、日本の心理学分野でも研究成果を認めてもらう実務上の手続きとして研究倫理制度が整備されてきたと指摘されている（河原・坂上2010, pp. iii-iv）。

上述の通り、心理学分野の研究倫理の倫理規程が整備され、倫理の「制度」が整備されてか

ら米国では40年以上が経過した。日本の同分野の代表的な学会である、日本心理学会が1991年に「倫理綱領」を策定し、わが国でも30年近くの実施の歴史がある。しかし、近年では研究への倫理規定の導入がもたらす弊害も指摘されており、その点は本稿第5節で言及する。

4. 研究倫理・評価倫理における「制度」の分析

心理学諸学会の倫理規程や日本評価学会の「評価倫理ガイドライン」は梅津のフレームワークの「制度」に含まれる。そこで本節では、まず心理学分野の倫理規程に記載された行動指針から研究の相対的な位置づけを明確にする。その上で、研究協力者の保護に関連する3点（倫理審査、インフォームド・コンセント、統制群³の設定）に着目し、「評価倫理ガイドライン」やAEAのGuiding Principlesと適宜比較し、評価分野への適用に際しての留意点を論考する。

（1）研究の位置づけ

各学会の倫理規程では、研究を実践活動との対比でどのように位置づけているだろうか。日本心理学会と日本教育心理学会では、研究は研究協力者に対してリスクの高い行為を含み、心理学の知見を利用する他の活動（教育、臨床活動）とは異なる扱いとしている。但し、倫理規程は研究と実践活動を明確に区分する基準を設定していない。従って、その判断は研究者自身、もしくは研究者の所属する機関に委ねる形式と考えられる。

表2に示す通り、日本心理学会（2009）の「倫理規程」では、大別して「研究と発表に関する倫理」と「社会における職務上の倫理」に対して行動指針を示す。前者は研究一般や研究分野別を、後者は心理学教育や臨床実践等の心理学を用いた諸活動を扱う。前者の研究一般及び全分野の行動指針では、倫理審査委員会による審査が義務付けられる一方、心理学の実践活動では審査への言及はない。倫理審査委員会は倫理上の問題を未然に防ぐ役割を果たすため、「倫理

表2 日本心理学会「倫理規程」の構成

はじめに
第1章 心理学にかかわる者の責任と義務
第2章 研究と発表における倫理
2.1 研究
2.1.1 研究一般
2.1.2 実験研究
2.1.3 調査研究
2.1.4 フィールド研究
2.1.5 臨床研究
2.1.6 動物を対象とした研究
2.2 発表
2.3 データベース
第3章 社会における職務上の倫理
3.1 教育・トレーニング
3.2 カウンセリングと介入
3.3 アセスメント
3.4 福祉の支援
3.5 司法、メディアとその他の社会的発言
第4章 倫理問題の解決

(出所) 日本心理学会 (2009) より筆者抜粋

規程」では、研究はそれ以外の活動に比べてリスクが高いと位置づけられる。具体例を挙げると、心理学実験には感覚刺激やだまし行為を伴うものがあり、フィールド研究では研究協力者のプライバシーに触れる機会が多い（日本心理学会 2009、pp.13-14、p.17）。

表3に示した日本教育心理学会（2000）の「倫理綱領」内の「倫理規定1. 人権の尊重」では「研究・実践活動の協力者」の尊厳への配慮が言及され、研究と実践活動が倫理規程の範囲と推察される。研究に関して、インフォームド・コンセントや研究中断を扱う「2. 研究実施のための配慮と制限」という項目がある。他方、研究以外の活動でも協力者の意志は尊重されるが、本人意思での中断は明示的に規定されず、研究では協力者の意志の尊重がより明確である。

研究と実践活動では対象者への配慮が異なるが、臨床研究では両者の線引きが曖昧になる。つまり、臨床研究では臨床活動の効果改善に繋がる手法が研究される一方、臨床活動では治療・

表3 日本教育心理学会「倫理綱領」の構成

・一般綱領
綱領1 人権および人間の尊厳に対する敬意
綱領2 学問上および専門職上の自覚と責任
・倫理規定
1. 人権の尊重
2. 研究実施のための配慮と制限
3. 情報の秘密保持の厳守
4. 公開に伴う責任
5. 研鑽の義務
6. 倫理の順守

(出所) 日本教育心理学会 (2000) より筆者抜粋

支援プロセスの一部として介入の効果が測定される。両方とも対象者の問題解決に繋がる介入と効果測定が行われるため、表面的な活動内容に基づく線引きが困難となる。

日本心理臨床学会（2016c）の「倫理綱領」（前文）では、会員は「その臨床活動及び研究によって得られた知識と技能を人々の心の健康増進のために用いるよう努める」ことを求められ、倫理規程はこの2つの行為を対象とする。また表4に示す、同学会が別途定めた「倫理基準」の第2条の2は、臨床業務は「学会水準で是認され得ない技法または不適切とみなされる技法を用いてはならない」と規定する（日本心理臨床学会 2016b）。加えて、同基準（第5条の1）では、研

表4 日本心理臨床学会「倫理基準」の構成

(責任) 第1条
(技能) 第2条
(査定技法) 第3条
(援助・介入技法) 第4条
(研究) 第5条
(秘密保持) 第6条
(公開) 第7条
(他専門職との関係) 第8条
(記録の保管) 第9条
(倫理の順守) 第10条

(出所) 日本心理臨床学会 (2016b) より筆者抜粋

究の行動指針として「対象者に対して通常の臨床活動以外の介入手続きを加える研究計画を立てる時は、研究の意義の検討、不利益の内容が道義的に認められる範囲内であることを確認する」と規定する。この記述は、研究自体を高リスクとはみなさず、研究に伴う介入が一般的な臨床活動ではない場合に、研究は倫理面の検討を要すると解釈できる。この区分方法は前述の「承認モデル」に依拠すると考えられる。

日本評価学会の「評価倫理ガイドライン」、AEAのGuiding Principles共に、研究と実践活動を区分した行動指針を設定していない（日本評価学会 2012、AEA 2018）。その理由として、いずれも外部・内部の評価委託者を前提とし、研究目的ではない評価業務を主に想定した倫理規程である点が挙げられる。この点で、評価分野の職業倫理は、実務家を念頭におくAPAの初期段階に近いと推察される。

(2) 倫理審査

倫理審査委員会は、客観的な立場から、研究毎にリスクの程度と対応を精査し、問題を防止する機能を担う。研究者が各研究で個別に配慮する分には倫理の「実践」に留まるが、倫理審査委員会が倫理的配慮を精査すると、研究実施に倫理の「制度」が介在することになる。

本稿で取り上げる学会のうち、日本心理学会の規程では、研究実施には倫理審査を行う委員会による承認が求められる。しかし、上記学会の倫理規程は、倫理審査の内容、つまりどの項目をどのような基準をもって審査するか、については具体的に記述していない。倫理審査の要否と同様に審査事項も、その判断は研究者自身、もしくは研究者の所属する機関に委ねられると推察される。

日本心理臨床学会（2016a）の「研究倫理審査委員会に関する運用内規」の第2条には、同会の委員会や専門部会、同会の指定機関等から申請された「研究計画の内容、計画の実行及びその成果の公表について審査を行う」ため、倫理審査委員会を設置する規程がある。この倫理審査で留意すべき事項として、第3条に4点（a. 個人情報の保護とデータの管理、b. 対象者の利益およ

び不利益、c. 心理臨床学としての社会貢献度、d. 対象者の理解と同意（インフォームド・コンセント））が挙げられている。倫理審査で特記された4項目は、研究の意義を問う一方、研究協力者のコストやリスクの負担とその対応に関連する。つまり、①研究の成果が十分あるか、②研究の費用やリスクが社会及び研究協力者に許容可能であるか、が審査される。注意すべき点は、臨床研究が効果の確立していない介入を伴うため、研究協力者は研究に係るコストやリスクを負担する一方、研究から受益できない可能性がある。そこで研究協力者に研究参加を強要する事態を防止するために、インフォームド・コンセントの手続きが倫理審査で事前に精査されている。

前述の通り、「評価倫理ガイドライン」、AEAのGuiding Principles共に、研究とそれ以外の活動を区分して行動指針を提示せず、倫理審査への言及はない。但し、評価協力者の安全、インフォームド・コンセント、プライバシーに関する情報の適切な管理は評価プロセスで配慮すべき項目としている。一方、評価の意義に関しては、「評価倫理ガイドライン」では評価を依頼する者に対し、業務開始前に効果的な評価結果の利用方法を検討すべき点を求めるが、評価自体の意義からその是非を判断するほど踏み込んだ精査は想定されていない。

評価分野で業務実施前に評価の意義に踏み込んだ検討がなされない一要因として、アクションリサーチのように評価実務と研究関心が一致する場合を例外として、大半の評価業務で研究とは異なる目的が設定される点が指摘できる。研究で重視される目的は一義的には知識の創造であり、すでに結論が定まったテーマに対し追加の研究は必要性が薄い。他方、評価では、知識の創造は目的の1つであるが、プログラムや政策の説明責任や改善も重要な目的となる（Patton 1997, p.51）。公的部門で一般に関心もたれるのは、後者（プログラムや政策の説明責任や改善）であろう。一般的に有効とされる介入であっても、特定の状況での有効性の精査に個別の評価を要し、また提言の導出には個々の介入のコンテキストを考慮する必要がある。同一のプログラム／政策に対し同一の評価目的かつ近似する

タイミングで評価を実施するような場合、評価を行う意義がなくなるが、そのような状況は一般的には発生しにくい。そのため、説明責任や改善を主な評価目的とする場合、評価する意義の有無ではなく、評価デザインから見た評価目的達成の可否や費用対効果といった別の視点から評価の意義を問うことになる。

制度運用上の課題として、審査制度がない場合の対処も検討を要する。日本心理学会（2009）の「倫理規程」の前文となる「はじめに」では、同規程が倫理審査を前提とする内容であり、倫理審査制度を持たない組織が少なくない点を承知しつつも、「将来的にはすべての組織が倫理委員会か、それに相当するものを持つことが望ましい」としている。一方、河原・坂上（2010）は、2009年11月に実施したアンケート調査で大学・研究機関に所属する回答者の4割が倫理審査制度を有しない機関に所属していた点に言及し、関係者に制度整備を働きかける一方、関係学会が審査を代行する例（日本公衆衛生学会、日本看護科学学会）に触れ、学会での審査を代替案として提示した（pp.160-165, pp.203-204）。前述の通り、日本心理臨床学会は、指定機関からの申請時に、倫理審査委員会を設置する制度があり、同委員会は所属研究機関内に倫理審査制度がない場合に利用可能なものと推察される。但し、同学会の各種規程では当該制度の趣旨を明示的に説明せず、この点は制度運用の実態に踏み込んだ精査を要する。

評価分野に関しては、AEAや日本評価学会の倫理規程では倫理審査委員会への具体的な言及はない。また、大半の評価者が論文投稿を前提としない評価業務に従事し、大学等に設置された倫理審査委員会を利用できない。

Cooksy（2005）は、米国の評価者に対し、倫理審査の要否の目安として、①論文作成や著作を通じて知識の創出を評価目的とする場合は研究とされ、倫理審査の対象となりうる、②省庁や公的機関からの補助金が評価予算となる場合、米国政府の規則（45CFR46）の対象となり、それに基づき審査要否が判断される、の2点を提示している（pp.354-355, p.357）。

他方、留意点として、評価分野で倫理審査を

要すると判断されても、評価に審査プロセスを新たに設ける必要はない状況も存在する。例を挙げると、国内公共事業やODA事業のように、評価とは別個に、介入自体が事前審査されるケースがあり、その場合には、事業評価の実施とは関係なく、介入によりネガティブな影響をうける層への配慮が審査される。

（3）インフォームド・コンセント

社会科学研究におけるインフォームド・コンセントは、研究協力者が研究による便益が本人ではなく、社会全般に還元されることを理解した上で、自発的に研究参加を承諾するプロセスとされる（藤本 2007、p.171）。さらには、研究協力者に対し、ネガティブ・リスクのある研究では、調査に先立ってインフォームド・コンセントを得ることが倫理面の問題を予防する措置となる。インフォームド・コンセント自体は研究倫理の「実践」であるが、「制度」である倫理規程ではその手順を規定する。

日本心理学会（2009）の「倫理規程」では動物を対象とする以外のすべての研究において、研究協力者に原則、文書で研究内容の理解を確認することが規定されている。また、代諾者が必要となるケースにも説明があり、子供、障害や疾患を有する人、外国人等、通常の方法では研究内容につき理解が難しい場合、代諾者からの書面同意が求められている。

同様に、日本教育心理学会（2000）の「倫理綱領」は、研究参加のインフォームド・コンセントに際し研究協力者の発達水準に応じた説明、必要に応じ保護者の代諾を規定している（「2. 研究実施のための配慮と制限」（2）、同（3））。

加えて、臨床におけるアセスメントであっても対象者からの同意が必須とされている。日本心理学会（2009）の「倫理規程」は、実践活動の心理アセスメントにもインフォームド・コンセントを求める。同じく、日本心理臨床学会（2016b）の「倫理基準」は、心理検査等に際して対象者、もしくは代諾者の同意を規定する（第3条）。

評価においても、評価協力者は協力のコストを負担するが、評価による改善は社会全体に裨

益が及ぶ。他方、評価に特有の留意点として、評価では集団が介入の対象となることが挙げられる (Oaks 2002, p.464)。介入に際し、受益者全員を対象とするインフォームド・コンセントは時間を要するプロセスであり、インフラ事業のようにその実施が困難な介入も存在する。

前述の通り、「評価倫理ガイドライン」は情報収集につき評価協力者に同意を得る点に言及している。しかし、リスクのある介入を伴う場合でも、評価実施に先立つインフォームド・コンセントは想定せず、自律的な意思決定が難しい層への配慮はなされていない。加えて、介入の受益者が幅広い場合のインフォームド・コンセントについてのガイダンスも提供されていない。

(4) 統制群の設定

介入研究において統制群の設定は必須だが、便益をもたらす可能性がある介入を特定集団に意図的に行わないことになる。統計群の設定は研究協力者間に不平等をもたらす恐れがあり、議論を呼びやすい研究手法といえる。

日本心理学会 (2009) の「倫理規程」(「2.1.5 臨床研究」「7.統制群の設定についての注意」)は、臨床研究で統制群を設定する場合、統制群に含まれる研究参加者に不利益となる可能性を指摘する。同倫理規程は、「統制群の設定は慎重に行い、可能なら別の方法を工夫して研究を行う」との方針であり、統制群の設定を極力避ける姿勢をとる。つまり、統制群に入る研究協力者の不利益を考え、統制群の設定は他の手段がない場合に限定されると解釈される。さらに議論を進めると、統制群の人々の相対的不利益が倫理上の基準と思考される。具体的には、介入の効果が確認されていない場合に、統制群が介入群に比べ不利益を被るとは見なせず、統制群の設定が可能となると解釈される。

評価が介入の割り付けに関与するケースが増加しており、統制群に入る人々の不利益が新たな倫理上の課題になりつつある。「評価倫理ガイドライン」に反映すべき事項として、倫理的に許容される割り付け手順を評価デザインに組み込むべき点を挙げたい。他方、注意が必要な点として、評価が介入の割り付けに関与しない状

況が存在する。例を挙げると、国内公共事業やODA事業では、評価が介入の形成や実施から切り離され、異なる部署や主体が評価と介入のそれぞれに責任を有する場合がある。従って、評価分野では、介入を割り付ける責任を誰が負うのかが倫理的な判断を左右する根拠となる。

しかしながら、「評価倫理ガイドライン」では、評価を目的とした介入の割り付け、統制群の設定を前提としないため、倫理上許容される割り付け手順や評価が割り付けに責任を有する場合とそうでない場合の対応に関して、評価者は行動指針を得ることができない。

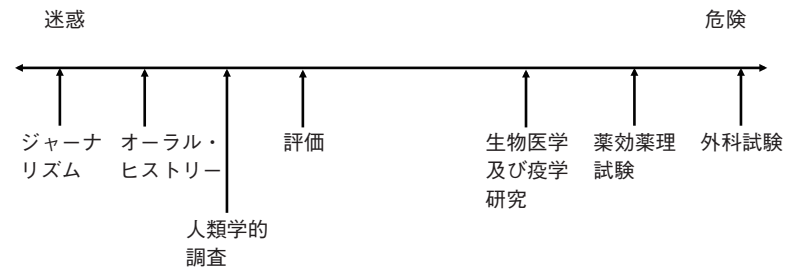
5. 研究倫理・評価倫理における「実践」から「制度」「理論」へのフィードバック

心理学では、比較的高いリスクのある研究手法と同一の基準を他手法に充てはめる弊害が指摘されている。心理学研究には、心理療法の効果測定も入り、図2で比較的「危険」とされる生物医学及び疫学研究に近い調査もある (Oaks 2002, p.449)。また、心理学にはだまし行為等のリスクを伴う調査手法を利用する研究もある。但し、同分野では研究協力者の負うリスクには幅があり、プライバシーの侵害、匿名性への脅威といった図2の「迷惑」のみに相当する研究も多い。

弊害の具体例として、心理学における研究倫理が研究手法の選定に影響を与えた可能性が指摘されている。社会心理学やパーソナリティ心理学において行動観察に基づく研究が1970年代から減少傾向にあり、倫理審査がその一要因となったとの批判がある (Baumeister et al 2007, pp.398-399)。

米国のThe Journal of Personality and Social Psychologyでの1966、1976、1986、1996、2006の各年における掲載論文のうち、行動を直接観察する研究は1976年には8割程度であったが、1990年代以降は2割程度に低下した。その原因として、倫理審査委員会が質問票調査に比べ行動観察を侵襲性が高いと見なし、倫理審査が困難となるため、研究者が行動観察を忌避する傾向がある

図2 分野毎のリスク



(出所) Oaks (2002) p.449より筆者邦訳

点が指摘されている (Baumeister et al 2007, p.400)。心理学分野で倫理審査が調査手法選定に過度な影響を与えているとの指摘からは、研究倫理を「実践」に移すにあたり、「制度」がしばしば過度な配慮を求めるため、研究自体を歪めてしまうリスクを示す。

評価はだまし行為や心理的外傷につながる情報収集を伴うものではないと一般的には考えられている (Newman and Brown 1996, p.167)。そのため、評価分野における協力者のネガティブ・リスクは、直接に心身を害するものではないと認識されていた (図2を参照)。介入の割り付けに関与する評価では、評価協力者のネガティブ・リスクは図2のより右側の「危険」に近づくが、それ以外の大半の評価においては、評価協力者のリスクは相対的に低いままである。

さらには、倫理面での配慮が研究協力者の負担となるケースも報告されている (田代 2014, pp.12-14)。例えば、臨床分野の質的研究でステイグマとなる病気の患者が研究協力者となる場合、書面同意が実名を知られたくない研究協力者には負担となる場合がある。

ベルモント・レポートの基本原則「人格の尊重」は、研究協力者の自発的な研究参加を求めている。上記の例での書面同意は、自発的な研究参加を確認する目的を持つ。他方、同レポートの基本原則「善行」は、研究協力者に害をなさないことを強調する。図1が示す通り、基本原則は「理論」に含まれるものである。上記の例は、倫理の「実践」の場で、「理論」が提示する2つの基本原則が対立するケースであり、そのような状況では適切な倫理的判断が困難となる。

評価においても協力者の保護を巡り基本原則が対立することが認識されている。Weiss (1997) は、評価に関与した人々を傷つけないこと、収集された情報を歪めないこと、の2つの義務が対立する点に言及している (p.430)。この対立は、「評価倫理ガイドライン」における「人々への敬意」と「公共の利益への責任」及び「体系的調査」が衝突することを意味する。Newman and Brown (1996) は、「人々への敬意」を含む評価の基本原則が異なる規範倫理学上の源泉 (功利主義、義務論等) を有している点に言及し、それぞれの長所や欠点も述べている (p.36, pp.52-54)⁴。しかしながら、「人々への敬意」と他の基本原則が衝突する局面における適切な倫理的判断を扱う論考はなされていない。加えて、AEAの Guiding Principlesは、その前文にて基本原則が対立することに言及するが、具体的な対処方法を提示しない (AEA 2018)。

6. 考察

(1) 「評価倫理ガイドライン」の不備への対応

第4節では「評価倫理ガイドライン」で不備な点を指摘した。そこで本節では不備への対応を論考したい。まず、客観的な立場から評価協力者のリスクの程度を判断し、配慮事項を精査するため、倫理審査の導入が適切と思料される一方、その要否の判断基準が議論の焦点となる。個々の評価業務において倫理審査の要否を判断する場合、研究倫理の「承認モデル」に基づく、効果発現が確定していない介入を伴う評価

が倫理審査の対象となるだろう。具体例として、評価がパイロット・プロジェクトの割り付けに関与するケースが挙げられる。但し、パイロット・プロジェクトでも実験的な要素には個々に濃淡があり、その濃淡を判断する基準も必要となる。

評価協力者の尊重以外に倫理審査で精査すべき事項としては、個々の評価業務の意義ではなく、評価デザインの適切さや費用対効果の精査が中心となるだろう。しかし、この審査を評価の利用者（評価委託者等）や評価の実施者（外部評価者、内部評価者）以外の第三者が行うことは適切ではないと思料される。費用対効果を踏まえて評価デザインを許容し、評価結果の活用を判断するのは一義的には評価の利用者であり、また評価実施にあたってのフィジビリティを判断するのは、評価結果の責任を負う評価の実施者である。

評価分野における「制度」のあるべき姿を議論する際に、「実践」上の課題ー倫理審査制度を利用できる評価者が少ないーを無視できない。そのため、「制度」の実効性を担保するには、評価者の有するリソースの制約に配慮した制度設計が求められる。具体的には、審査対象となる評価業務の適切な線引き、倫理審査へのアクセス改善が要検討の課題となるだろう。

リソースの節減の観点から、評価分野では既存の審査手続きの利用も選択肢に入るだろう。心理学研究では、研究自体が介入の主体となるため、研究の審査では介入の是非も問われる。他方、評価分野では事業で介入を行う際に、評価とは別個に、事前の審査でプロジェクトの負のインパクトと被影響者が特定され、その影響への対処方針が精査される。従って、評価に特有の事項（例えば、評価協力者へのインフォームド・コンセント、効果測定に向けた介入の割り付け）がこの審査で適切に検討されれば、評価協力者への配慮を担保できる実質的な倫理審査となる。

次いでインフォームド・コンセントについては、「評価ガイドライン」において評価が介入の割り付けに関与する状況を想定した行動指針が望まれる。さらには、インフォームド・コンセ

ント全般につき、評価協力者が評価内容を理解できる能力に応じた説明の必要性、時には代諾者による同意も配慮すべき事項となる。

インフォームド・コンセントの留意点として、対象範囲の慎重な精査を挙げたい。受益者全員を対象とするインフォームド・コンセントは時間的な制約から困難が生じ、大規模インフラ事業のように受益者を特定し、全員の同意が難しい状況も存在する。評価に伴う介入に先立ちインフォームド・コンセントを行い、評価協力者の同意を求めるのであれば、適切な範囲の設定が今後、議論すべき点として浮上する。

加えて、統制群の設定を規定する行動指針も今後、検討が必要になる。留意点として、評価が介入の割り付けに関与しない場合が挙げられる。その場合、介入の割り付けに関する倫理面の課題から、評価を中止する判断は適切であろうか。むしろ課題があるからこそ、評価すべきという議論も成り立つ。心理学研究で割り付けに倫理面の課題がある場合、研究者に割り付けの責任があるが、評価分野では割り付けの責任を常に評価者に帰することはできない。

(2) フィードバックに関する考察

第5節で言及した「実践」からの「制度」「理論」へのフィードバックから、評価分野でも注意すべき点を提示したい。第1のフィードバックは、心理学分野での倫理審査の導入により、一部の研究手法が忌避されている点である。介入の割り付けに関与する評価では評価協力者のリスクは増える一方、それ以外の評価では研究協力者の負うネガティブ・リスクは従前と変わらない。評価業務毎のネガティブ・リスクの高低を踏まえず、厳しい基準の適用範囲を広げることが評価手法の選定を歪める弊害をもたらすだろう。

これまで評価では手法の依拠するパラダイムが多様であるため、強制力を持つ倫理規程を策定しにくい点が指摘されている（Fitzpatrick 1999, p.12）。この状況は無視し、「制度」が一部の評価手法を著しく困難にする規程を設け、その遵守を強いることは、評価手法の多様性を損なう。評価結果の有用性を高めるには、評価目的に合

致する評価手法を選定する方が適切であり、倫理審査の容易さから評価手法が選定されることは望ましくない。

もう1つのフィードバックは、「実践」において「理論」が提示する基本原則の対立が生じており、その結果、研究協力者への負担が大きくなる可能性がある。管見の限り、評価倫理の「理論」においても、「人々への敬意」と他の基本原則が衝突する具体的な局面を想定し、個々のケースにおける適切な対処方法は今後の論考を必要とする点と考えられる。加えて、評価倫理の「実践」において、評価者や倫理審査委員会メンバーが基本原則の衝突する局面を扱うケース教育を受け、倫理上の判断を経験することも有意義と思料される。

7. おわりに

評価に伴う情報収集や介入の割り付けが評価協力者の過度な負担やリスクに繋がる可能性は常にあり、「人々への敬意」の観点から適切な配慮を行う必要性は論を待たない。とりわけ評価が介入の割り付けに関与する状況は現行の「評価倫理ガイドライン」が想定するものではなく、評価倫理の「制度」を拡充すべき事項となる。

心理学分野では、研究は実践活動よりもリスクが高いと位置づけられ、日本心理学会の倫理規定では研究は倫理審査を必須とする。倫理審査は、研究協力者が被るリスクの程度とその対応を客観的な立場から精査する「制度」である。同様に、評価協力者に著しいリスクが生じる可能性がある場合、類似する審査制度が望まれ、その可否の判断基準を設定する必要がある。なお、判断基準の設定に際しては、研究倫理の「承認モデル」が参考となるだろう。加えて、倫理審査のリソースの節減に向けた既存の審査手続きの活用、加えて幅広い層への倫理審査へのアクセスを可能にする制度設計が取り組むべき課題として浮上した。

「評価倫理ガイドライン」は情報収集に関するインフォームド・コンセントを求めるが、評価が介入の割り付けに関与する状況でもインフォ

ームド・コンセントが必要と思料される。さらに、評価が介入を伴う場合、インフォームド・コンセントの範囲設定が検討を要する点として明らかとなった。また、研究倫理との比較から、評価協力者が説明を理解できる能力に応じた同意方法も検討すべき事項と思料される。

さらには、許容される統制群の設定も「評価倫理ガイドライン」で扱うべき事項であろう。但し、評価が介入の割り付けに責任を有する場合とそうでない場合があり、両ケースの混同を避ける指針も併せて検討されるべき事項となる。

研究倫理の「実践」から憂慮すべき点として、倫理審査が一部の研究手法を排除する可能性が示唆された。さらに、研究倫理の「実践」にて「理論」が提示する基本原則の対立が生じ、その結果、倫理面での配慮が研究協力者の負担を大きくするケースも明らかとなった。いずれの課題も評価分野において生じる可能性がある。

本稿は、心理学諸学会の倫理規程から評価倫理の「制度」にて留意すべき点を指摘したが、その多くは問題提起に留まる。また、その留意点も網羅的、体系的に導出されておらず、個別事例を心理学全般の事象として過度に一般化している可能性がある。加えて、本稿の論考は文献に頼るが、学会、大学、研究機関における研究倫理の実際の制度運用から導出できる示唆も多いだろう。

注記

- 1 本稿における見解は、筆者個人としてのものであり、筆者が所属する組織の見解とは必ずしも一致しない。
- 2 本稿はベルmont・レポートの邦訳として、津谷喜一郎ら訳（2001）「ベルmont・レポート」臨床評価刊行会編『臨床評価』28（3）、559-568に依拠した。
- 3 「統制群」は主に心理学及び教育学研究で使用される用語であり、保健医療ほかでの「対照群」、「コントロール群」と同義である。
- 4 規範倫理学は倫理学の一分野であり、人々のあるべき行動の基礎となる「規範」(norm)を探求し、道徳的な善さや正しさを扱う。

参考文献

- 梅津光広 (2002) 『ビジネスの倫理学』、丸善出版
- 川瀬貴之 (2013) 「ベルモント・レポートに見る臨床研究の倫理指針の問題点」、石井徹哉編『千葉大学大学院人文社会科学研究所研究プロジェクト報告書』、258：61-70
- 河原純一郎・坂上貴之編著 (2010) 『心理学の実験倫理：「被験者」実験の現状と展望』、勁草書房
- 笹栗俊之・池松秀之編 (2011) 『臨床研究のための倫理審査ハンドブック』、丸善出版
- 田代志門 (2011) 『研究倫理とは何か：臨床医学研究と生命倫理』、勁草書房
- 田代志門 (2014) 「フィールドワーカーが倫理審査に出会うとき-医療現場における質的研究の「困難」、『質的心理学フォーラム』、6：9-17
- 日本教育心理学会 (2000) 『日本教育心理学会倫理綱領』、<https://www.edupsyh.jp/wpcontent/uploads/2017/04/4ad9a1d91eb958c280f32a0600fd8b311.pdf> (2019年1月14日アクセス)
- 日本心理学会 (2009) 『日本心理学会倫理規程』、http://www.psych.or.jp/publication/inst/rinri_kitei.pdf (2019年1月14日アクセス)
- 日本心理臨床学会 (2016a) 『研究倫理審査委員会に関する運用内規』、http://www.ajcp.info/pdf/rules/017_rules_514.pdf (2019年1月14日アクセス)
- 日本心理臨床学会 (2016b) 『倫理基準』、http://www.ajcp.info/pdf/rules/015_rules_512.pdf (2019年1月14日アクセス)
- 日本心理臨床学会 (2016c) 『倫理綱領』、http://www.ajcp.info/pdf/rules/014_rules_511.pdf (2019年1月14日アクセス)
- 日本評価学会 (2012) 『評価倫理ガイドライン』、http://evaluationjp.org/files/JES_Guidelines_for_the_Ethical_Conduct_of_Evaluations20121201.pdf (2019年1月14日アクセス)
- 八田進二・町田祥弘 (2003) 「会計プロフェッションの自主規制における職業倫理の位置づけ」、『企業会計』、55 (3)：39-48
- 藤本加代 (2007) 「アメリカ合衆国における「IRB制度」の構造的特徴と問題点ー日本の社会科学研究における研究対象者保護制度の構築に向けてー」、『先端社会研究』、6：165-188
- AEA (2018). Guiding Principles for Evaluators .<https://www.eval.org/p/cm/ld/fid=51> (2019年1月14日アクセス)
- Baumeister, R.F. et al (2007). Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior? *Perspectives on Psychological Science*, 2(4), 396-403.
- Cooksy, L. J. (Ed.) (2005). Ethical Challenges: The Complexity of the IRB Process: Some of the Things You Wanted to Know About IRBs but Were Afraid to Ask. *American Journal of Evaluation*, 26(3), 352-361.
- Fitzpatrick, J. L. (1999). Ethics in Disciplines and Professions Related to Evaluation. In J. L. Fitzpatrick and M. Morris (Eds.), *Current and Emerging Ethical Challenges in Evaluation*. San Francisco, CA: Jossey-Bass, 5-14.
- Newman, D.L. and Brown, R.D. (1996). *Applied Ethics for Program Evaluation*. Thousand Oaks, CA: SAGE Publications
- Oaks, J.M. (2002). Risks and Wrongs in Social Science Research: An Evaluator's Guide to the IRB, *Evaluation Review*, 26(5), 443-479.
- Patton, M.Q. (1997). *Utilization-Focused Evaluation*(3rd Edition) . SAGE. (大森彌監修、山本泰・長尾眞文編 (2001) 『実用重視の事業評価入門』、清水弘文堂書房)
- Sabourin, M. (1999). The Evolution of Ethical Standards in the Practice of Psychology: A Reflection on APA Code of Ethics. *The Japanese Journal of Psychology*, 70(1), 51-64. (中島定彦訳 (1999) 「心理学における倫理基準の発展-アメリカ心理学会倫理規定の一省察-」、『心理学研究』、70 (1)：51-64)
- Steneck, N.H. (2006). Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions. *Science and Engineering Ethics*, 12 (1), 53-74.
- Weiss, C.H. (1997). *Evaluation* (2nd Edition). Prentice Hall (佐々木亮監修、前川美湖・池田満監訳 (2014) 『入門 評価学 政策・プログラム研究の方法』、日本評論社)

(2019.9.19 受理)

Considerations on Respect for Evaluation Participants in Evaluation Ethics: Implications from Research Ethics in Psychology

Nobuyuki Kobayashi

OPMAC Corporation
kobayashi@opmac.co.jp

Abstract

This paper discusses respect for people, an important principle of evaluation ethics, from the viewpoint of ethical considerations for research participants in Psychology. Based on the discussion, this paper withdraws the implications for issues to be reflected in the Guidelines for the Ethical Conduct of Evaluations, which was established by the Japan Evaluation Society. As psychological research involves a higher risk than practical activities, an ethical review is required prior to research. When evaluation participants are to be exposed to a significant risk, a similar review is desirable. Thus, judgment criteria on the necessity of ethical review need to be researched. Since the Guidelines does not assume the situation that an evaluation decides assignment of intervention, informed consent in such situation and acceptable formation of control group also need to be discussed. In addition, some issues emerged from the practice of research ethics. Ethical review affected the selection of research method and researchers tended to avoid a certain type of research method. The ethical principles conflicted each other and this situation made ethical judgement difficult. As a result, the ethical considerations overburdened research participants in the case.

Keywords

evaluation ethics, research ethics, ethical review, informed consent, control group

【研究ノート】

庁が実施する評価の現状と課題

鏡 圭佑

総務省行政管理局 行政手続室

k.kagami@soumu.go.jp

要 約

本稿の目的は、庁が果たす機能の観点から、庁が実施する現行の評価制度における研究課題を提示することである。庁とは、中央政府に置かれる行政組織の形態の1つである。政策過程において果たす機能の違いから、庁は実施庁と政策庁に大別される。それぞれの庁の機能を明確にすることで、その庁の機能に適した政策評価のモデルを構想できる。本稿では、この政策評価のモデルと現行の評価制度の比較を通じて、評価制度における研究課題を抽出する。

今後の研究課題として、実施庁における学習型評価の在り方を検討する研究、実施庁における実績評価制度を特定の行政組織が所管する必要性を考察する研究、政策庁が実施する目標管理型政策評価および総合評価の学習に対する貢献度を把握する研究を提示する。

キーワード

実施庁、政策庁、実績評価、目標管理型政策評価、総合評価

1. はじめに

本稿では、庁が果たす機能の観点から、庁が実施する現行の評価制度における研究課題を提示する。

庁とは、中央政府に置かれる行政組織の1つである。庁は国家行政組織法（昭和23年法律第120号）第3条第2項および内閣府設置法（平成11年法律第89号）第49条第1項に規定があるように、府、省および委員会と同様に、その設置および廃止が法律によって決められる基幹的な行政組織である。

行政改革会議（以下「行革会議」という）は改革を進めるために、庁を実施庁および政策庁に分類した。実施庁は主に政策の実施を担う庁

であり、政策庁は主に政策の企画立案を担う庁である。本稿では、行革会議による類型に基づいて、実施庁および政策庁の機能を明確にする。

庁の機能を明確にすると、その庁に適した政策評価のモデルを構想できる。ここでは、政策評価を行政組織の活動の調査を通じて、その活動を改善するために必要な情報を産出する活動であると定義する¹。とくに、本稿では2つの庁の機能に着目して、庁が自らの機能をより適切に果たすために、その活動を改善する際に有益となる政策評価のモデルを提示する。

他方で、現行の評価制度のもとで、2つの庁は政策評価を実践している²。この研究では、実施庁の実績評価、政策庁の目標管理型政策評価および総合評価をとりあげる。

「庁の機能に適した政策評価のモデル」と「現行の評価制度」との比較を通じて、庁の政策評価における研究課題を抽出する。その際の中心視点は、評価学においても言及されてきた政策評価における統制および学習の観点である（山谷 2012、pp.220-225）³。ここでいう、統制とは上位の組織が下位の組織による機能の遂行を監視するための評価であり、学習とはその組織が自らの機能を改善するために行う評価を指す。2つの観点に基づいて、実施庁の評価制度における学習型の評価の必要性および政策庁における現行の評価制度の学習にとっての有用さの把握を今後の研究課題として提示する。なお、本稿で検討するモデルと評価制度は表1のようにまとめられる。

本稿には、行政組織と政策評価の関係をめぐる視座から研究を進める点に意義がある。管見の限り、特定の行政組織の機能に着目し、政策評価の課題を論じる研究は公表されてこなかった。代わりに、評価学では政策の性質と評価の関係を論じるアプローチが主流であるように考えられる⁴。ここでは、より根本的である行政組織の機能の観点から研究を進めることで、評価学における研究範囲の拡大に貢献するような視座の確立を試みる。

研究の構成は以下のとおりである。2.では庁を簡潔に定義し、実施庁および政策庁の類型を示す。3.ではそれぞれの庁の機能に適した政策評価を構想し、現行の評価制度を記述する。4.では庁の機能に適した政策評価と現行の評価制度を比較し、2つの庁の政策評価における研究課題を提示する。

表1 本稿で検討するモデルと評価制度

	政策評価のモデル	現行の評価制度
実施庁	省の長による統制のための評価 実施庁による活動改善のための評価	実績評価
政策庁	政策目標の見直しのための評価 政策手段の有効性の確認のための評価	目標管理型政策評価 総合評価

(出所) 筆者作成

2. 庁の定義と類型

ここでは、行政組織としての庁の特徴を明確にし、実施庁および政策庁の類型を説明する。

(1) 庁の定義

庁は国家行政組織法第3条第2項において、その設置および改廃が法律によって決められる組織であり、委員会と同様に、府省の外局として設置される。外局とは、府省に置かれる官房、局、部、課および室といった内部部局の対になる概念である（宇賀 2019、p.189）。国家行政組織法第3条第3項では省が内閣の統括のもとで内閣の行政事務を分担管理する行政組織であり、委員会および庁は省の外局であると整理されている。また、内閣府設置法第49条第1項においても庁は府の外局であると整理されている。外局である庁は内部部局と同様に、所属する府省から一定の統制を受ける。他方で、内部部局とは異なり、庁は一定の権限を自律的に行使する。

庁の長官には、行政組織に関する法制度に基づいて一定の権限が付与されている。まず、国家公務員法（昭和22年法律第120号）第55条によって、職員の任命権が付与されている。つぎに、国家行政組織法第13条第1項に基づいて庁の長官は政令および省令以外の規則あるいは特別の命令を制定でき、同法第14条第1項によって自らの所掌事務の範囲内で告示を発出できる。そして、同法第15条に基づいて庁の長官が自らの任務を実現するために、ほかの行政組織に資料および説明を求め、その行政組織の政策に意見を述べる権限が認められている。なお、内閣府に置かれる庁の長官も内閣府設置法第58条に基づいて国家行政組織法に基づいて設置される庁の長官と同様の権限が付与されている。

庁には、長官を補助するための内部部局が置かれる。国家行政組織法第7条第3項に基づいて、省の外局である庁には官房および部を、同法第7条第5項に基づいて庁、官房または部には課および室を置くことができる。内閣府に置かれる庁に関しても、内閣府設置法第53条第1項に基づいて官房および部の設置が可能である。さらに、同法第53条に基づいて庁、官房および部には課

および室を置くことができる。なお、同法第53条第2項によって、法律で特命担当大臣を置くことと決められている庁は局を設置することが可能である⁵。これらの内部部局に加えて、審議会、施設等機関、特別の機関⁶、地方支分部局といった附属機関を設置する庁もある。より具体的な組織構造は、設置法以降の規範によって定められる。

内部部局の活動を通じて、庁は政策を管理する。政策とは社会的な問題に対する解決の方針と手段である（秋吉・伊藤・北山 2015、p.4）。庁は所属する府省から委譲された所掌事務を実現するために、さまざまな政策を管理する。たとえば、国税庁は財務省設置法（平成11年法律第95号）第19条および第20条に基づいて、財務省の所掌事務のなかから、国税の徴収事業に関する事務を委譲されている。また、スポーツ庁は文部科学省設置法（平成11年法律第96号）第16条の規定により、文部科学省の所掌事務のうち、スポーツの振興に関する施策の総合的な推進等に関する所掌事務を委譲されている。

(2) 実施庁および政策庁

庁は政策過程において果たす機能の違いから、実施庁および政策庁に分類できる。前出の国税庁は実施庁に、スポーツ庁は政策庁に分類される。この類型を構築した行革会議は政策庁と実施庁を簡潔に定義し、中央省庁再編後に設置される庁を2つの類型から整理した。しかし、行革会議の最終報告において、それぞれの庁の機能は簡潔に定義されるのみである。行革会議による実施庁および政策庁に関する提言を制度に反映させるために、中央省庁等改革基本法が制定され、国家行政組織法が改正されたが、それらの法律にも詳細な定義はない。したがって、ここでは今後の研究に必要な範囲で実施庁および政策庁の機能を定義する。

なお、国家行政組織法別表第2において実施庁に該当する庁が列記されている。この整理に従うと、表2のように、現存する庁を分類できる。

① 実施庁

実施庁は主として政策の実施に関する機能を

表2 現存する庁の一覧とその類型

府省	庁	類型
内閣府	金融庁	政策庁
	消費者庁	政策庁
総務省	消防庁	政策庁
法務省	公安調査庁	実施庁
	出入国在留管理庁	政策庁
外務省	なし	なし
財務省	国税庁	実施庁
文部科学省	スポーツ庁	政策庁
	文化庁	政策庁
厚生労働省	なし	なし
農林水産省	林野庁	政策庁
	水産庁	政策庁
経済産業省	資源エネルギー庁	政策庁
	特許庁	実施庁
	中小企業庁	政策庁
国土交通省	観光庁	政策庁
	気象庁	実施庁
	海上保安庁	実施庁
環境省	なし	なし
防衛省	防衛装備庁	政策庁

（出所）筆者作成

果たす。より具体的には、実施庁の機能は政策実施の過程における個別具体的な事業の実施に求められる。行革会議は自らが目指す改革の基本方針である政策の企画立案と実施の分離に基づいて、政策庁と実施庁を分類し、実施庁が自律的かつ効率的に自らの事業を実施する体制を提言した（行政改革会議事務局OB会編 1998、p.73）。提言を受けて、国家行政組織法第7条第6項において実施庁の内部部局編成における一定の弾力性を確保する規定等が整備された。

現存する実施庁は、国家行政組織法別表第2に掲げられているとおり、公安調査庁、国税庁、特許庁、気象庁および海上保安庁である。これらの庁を概観すると、実施庁の事業には3つの特徴がある。すなわち、①作業量が膨大な事業（国

税庁および特許庁)、②秩序維持のための事業(公安調査庁および海上保安庁)、③広く国民が知る必要のある情報を提供する事業(気象庁)という特徴がある。

実施庁は、実施担当組織の活動を通じて事業を実施する。実施担当組織とは、大規模な人員をもって、実施庁の任務を直接的に実施する内部部局または地方支分部局等を指す。たとえば、国税庁には国税の徴収事業を遂行するために、2019年8月時点で11の国税局、518の税務署が置かれている。また、海上保安庁には海上の治安を維持するために、2019年8月時点で11の管区海上保安本部が置かれ、各管区海上保安本部には海上保安(監)部および海上保安署が設置されている。

実施庁が自らの事業を実施する過程は以下の2つの段階に大別される。

第1に、実施庁の目標を策定する段階がある。実施庁の目標は所属する省の長が決める。具体的には、中央省庁等改革基本法第16条第6項第2号では、実施庁が所属する省の長が実施庁の目標、実施の基準、その他の実施に必要な準則を決定すると規定されている。ここから、実施庁の活動に対する省の長による統制の存在を確認できる。政策庁とは異なり、実施庁には自らの目標を自らで策定する権限がない⁷。

第2に、実施庁が実施担当組織の活動を通じて、事業を実施する段階がある。実施庁は、行革会議において専門的な知見をもって政策を実施する組織として構想された。具体的には、「公正、中立、透明な行政執行と効率的で国民のニーズへの即応を重視した行政サービスの提供」が実施庁または独立行政法人を含めた政策実施部門に共通の使命とされた(行政改革会議事務局OB会編 1998、p.57)。すなわち、実施庁には、省の長から与えられた目標を実現しつつ、それにかかる金銭、人員、時間といったコストの抑制およびサービスの質の向上を実現するような実施が求められている。

② 政策庁

政策庁は政策目標を設定し、その実現のために、政策手段をもって政策対象に働きかける。

行革会議は省の外局である庁は実施庁のみを残し、政策庁は例外として設置を認めると考えていた。例外としての政策庁は、①その庁の政策立案作業が所属する省の内部部局が担う他の作業から明確に区別される一定のまとまりをもつこと、②その政策立案を内部部局ではなく外局に担わせる特段の必要性があること、③担当する事務および事業の独立行政法人化または業務の大幅な縮小が行われるものではないことという3つの条件すべてを満たす場合に認められた(行政改革会議事務局OB会編 1998、pp.72-73)。

現存する政策庁の数は、12である。具体的には、金融庁、消費者庁、消防庁、出入国在留管理庁、スポーツ庁、文化庁、林野庁、水産庁、資源エネルギー庁、中小企業庁、観光庁および防衛装備庁である。行革会議では例外として位置づけられていたが、行革会議の提言が出された後も、後述するように、内閣の政治的な判断に基づいて政策庁が定期的に設置されている。

政策庁は政策目標の設定および政策手段の適用を通じて自らの機能を果たす。政策庁は所属する省から委譲された特定の領域における予算案または法律案等の政策案を立案する。政策案において、政策庁の政策目標、政策対象および政策手段が規定される。政策対象とは、政策庁が政策目標を実現するために働きかける個人または集団である。また、政策手段とは、政策庁が政策対象の行動を政策目標にとって望ましい方向に向けさせる手段である。山谷清志(2012a、p.41)によれば、政策手段は分配、再分配、規制と規制緩和、広報・PR・教育、税制および制度の創設と改変に分類できる。たとえば、防衛装備庁は防衛装備品等の開発および生産基盤の強化のために、企業、大学、または研究機関等の政策対象に補助金等の政策手段を適用する⁸。

政策庁が自らの機能を果たす過程は、以下の2つの段階に大別される。

第1に、政策を企画立案する段階がある。林野政策、中小企業政策、スポーツ政策または防衛装備政策を一瞥しても分かるように、政策課題、政策目標、政策対象および政策手段は政策領域ごとに多様である。政策庁は自らが取り組む政策課題を理解するために必要な専門性を備えた

うえで、実現可能な政策目標と有効性がある政策手段の構想といった創造的な作業を行わなければならない。

第2に、政策庁は政策対象に政策手段を適用する⁹。政策目標は、政策対象の望ましい行動によってはじめて実現する。したがって、政策庁は政策手段を通じて、政策対象の行動を政策目標に合致するように変化させる必要がある¹⁰。政策対象は個人または組織ごとに異なる価値観を有している以上、政策庁は予測に基づいて多様な政策対象の内面を有効に変化させるように政策手段を適用しなければならない。

3. 庁の機能と政策評価

ここでは、2.で示した実施庁および政策庁の機能の観点から庁に適した政策評価のモデルを構想し、実施庁の実績評価、政策庁の目標管理型政策評価および総合評価を整理する。

(1) 実施庁と政策評価

まず、実施庁の機能に適した政策評価のモデルを構想し、現行の評価制度を整理する。

① 実施庁の機能に適した政策評価

上述のように、実施庁は2つの段階を通じて機能を果たす。第1に、省の長が実施庁の目標を決定する段階があった。第2に、実施庁が実施担当組織の活動を通じて目標の実現に努める段階があった。2つの段階ごとに実施庁に適した政策評価のモデルを構想できる。第1の段階に対応する評価（以下「第1の評価」という）は省の長による実施庁の統制を支援するための政策評価であり、第2の段階に対応する評価（以下「第2の評価」という）は実施庁による自らの活動を改善するための政策評価である。

2つの政策評価のモデルは、本稿の冒頭で言及した評価学における統制と学習の観点から導出されている。すなわち、第1の評価は統制を志向する評価の考え方から派生している。他方で、第2の評価は学習を志向する評価の考え方に基づいている。

第1の評価では、政策評価を通じて、省の長による実施庁の統制が試みられる。上述のように、省の長は実施庁に達成すべき目標を提示する。省の長は、実施庁が目標のとおりに事業を遂行しているか否かを監視しなければならない。仮に、実施庁が目標を達成できていない、または目標から逸脱した行動をしていた場合には、統制を加える必要がある。第1の評価は、こうした監視または統制に必要な情報を生み出すツールになる。必要になる情報は、省の長が提示した目標に対応した実施庁の活動実績を表す情報である。

第2の評価では、政策評価を通じて、実施庁による自らの活動の改善が試みられる。実施庁は所属する省の長が提示した目標に従いつつも、事業を遂行する過程においては自律的に活動する。こうした活動における実施庁の判断が、事業の成否または成果の向上にとって重要となる。第2の評価は、実施庁が自らの判断をよりよくするための学習のツールになる。第2の評価によって生み出すべき情報は、実施庁の活動による政策対象または社会の変化に関する情報¹¹、実施庁の創意工夫における失敗例または成功例といった活動改善の参考になる情報が考えられる。

2つの政策評価には、いくつかの重要な相違点がある。たとえば、行為主体としての実施庁の捉え方における違いがある。すなわち、第1の評価では実施庁が所属する省の長に従属する側面が強調され、第2の評価では実施庁が自律的に事業を遂行する側面が強調される。また、第1の評価では評価の際に省の長が事前に提示した目標が参照され、第2の評価では実施庁の活動の後に産出された成果が参照される。さらに、第1の評価の目的は統制にあり、第2の評価の目的は学習にある。これらの視点の違いおよび異なる目的を考慮すると、単一の評価制度のなかで2つの評価を同時に行うことは困難であり、評価制度を別個に設計する必要がある。

② 現行の実施庁の評価制度

現行の評価制度のもとで、実施庁の活動を評価するために、実績評価が実践されている。中央省庁等改革基本法第16条第6項第2号では、省

の長は実施庁の活動の目標を提示するだけでなく、その活動を評価しなければならないと規定されている。この規定に基づく政策評価が実績評価である。実績評価を通じて、省の長が提示した目標を実施庁がどの程度まで実現しているのが毎年度測定されている。

実績評価は省の長が提示した目標と実施庁の活動の実績を照合する形で行われる。省の長は、年度ごとに、実施庁に活動の目標を提示する。たとえば、実施庁が処理する個々の案件の処理の迅速化、処理件数の増大、対象者の満足度の向上等が目標として提示されている。中央省庁等改革推進本部が定めた「中央省庁等改革の推進に関する方針」に従って、原則的に、目標は定量的に測定可能な指標として提示される（中央省庁等改革推進本部 1999）。これらの指標に基づいて、毎年度の実施庁における実施担当組織の活動の実績が測定される。そして、省の長の示した目標と実施庁の活動の実績値が照合される。その結果に基づいて、省の長は実施庁の活動が妥当であったか否かを判定し、場合によっては必要な措置をとる。

実績評価では、アウトプットが測定される。アウトプットとは、行政組織に投入された人員、資金、時間等の資源を用いた目標実現のための活動量を意味する。すなわち、インプットをアウトカムに変換する活動である。ここで、インプットとは投入される資源であり、アウトカムとは政策の実施を通じて社会にもたらされた成果である。省の長の目標のほとんどは、実施庁によるアウトプットの経済性、効率性、迅速性といった観点から設定されている（東 2002、pp.206-207）。アウトプットの評価によって、省の長は目標において示した期待を実施庁が実現したか否かを把握する。他方で、実施庁の活動による政策対象や社会の変化はアウトカムに該当するため、基本的に、実績評価では把握できない点に留意する必要がある。

評価期間は、毎年度である。この期間の設定によって、実績評価は実施庁による事業の進行政管理および省の長による統制のツールとして機能するようになる。毎年度の実施庁の活動測定を通じて、制度上は省の長は自らが提示した目

標と実施庁の活動との一致をつぶさに確認できる。さらに、実施庁の活動における不適切な点を直ちに把握し、必要な措置を迅速に講じることが可能になる。また、定期的な記録の蓄積によって、省の長が実施庁の平均的なパフォーマンスを学習し、より適切な目標を設定できるようになる。

以上の諸特徴から判断すると、実績評価では上述の第1の評価のみが志向されている。山谷が述べている通り、統制を志向する評価と学習を志向する評価は両立しがたいジレンマの関係にある（山谷 2012、p.223）。さらに、上述のように、単一の評価制度において第1の評価と第2の評価を両立することは困難である。実際に、平成30年度に実施庁（公安調査庁、国税庁、特許庁、気象庁、海上保安庁）が実施した事業に対する実績評価の報告書では、省の長の提示した目標、活動の実績値、照合の結果が簡潔に記載されているのみである¹²。

（2）政策庁と政策評価

つぎに、政策庁の機能に適した政策評価のモデルの構想および現行の評価制度の概観を行う。

① 政策庁の機能に適した政策評価

政策庁は、政策目標および政策手段を含む政策を企画立案する。政策の企画立案という機能から、政策庁に適した2つの政策評価のモデルを構想できる。すなわち、政策目標を見直すための政策評価および政策手段の有効性を確認するための政策評価があげられる。

2つの政策評価はともに、本稿の冒頭において述べた学習を志向する政策評価である。政策庁に学習型の評価が求められる理由は、政策庁の機能の性質および機能を果たす環境から説明できる。政策の企画立案における政策目標の設定および政策手段の選択は「政策デザイン」とも呼ばれるように、広範な裁量に基づく、きわめて創造的な性質の活動である（足立 2009、pp.23-28）。また、政策の企画立案は、社会経済環境および政策対象の行動といった不確実性が高い事象に対する予測に基づく情報を利用して行われる。創造性および不確実性が高い政策の企画立

案であるからこそ、政策庁は2つの政策評価を用いて、自らの判断に基づく政策によって社会がどのように変化したのかを調査し、調査結果に基づく学習を通じて、政策目標および政策手段に不断の改善を加える必要がある。

第1に、政策庁が企画立案した政策目標に対する政策評価がある。政策庁は自らの判断によって、当該政策領域における政策目標を設定する。政策目標は、政策によって実現すべき政策対象の望ましい状態である。政策目標が妥当であるか否かは、政策を実施しなければ判断できない。仮に、当初は妥当であっても、後の社会経済環境の変化によって、政策目標が妥当ではなくなる可能性がある。したがって、政策庁自身が政策評価を通じて、政策目標の妥当性を確認する必要がある。第1の評価では、政策目標が実現可能であるのか、政策目標は政策対象または社会の実態に照らして妥当であるのか、そもそも政府が政策目標掲げて政策介入をする必要があったのか調査される。

第2に、政策庁が政策対象に適用する政策手段の評価がある。政策目標を実現するためには、政策庁が適切な政策手段を選択し、政策対象の行動を確実に変化させるように、それを適用する必要がある。政策庁は政策対象の行動パターンおよび社会経済環境等の予想に基づいて、政策手段の設計および適用を行う。政策庁はこの予想のとおり政策手段が機能しているのか否か、政策手段にどの程度の有効性があるのかを確認する必要がある。さらに、確認の結果に基づいて政策手段を改善することも求められる。

② 現行の政策庁の評価制度

ここでは、目標管理型政策評価および総合評価を検討する。

目標管理型政策評価とは、総務省によって整備された実績評価を体系的かつ標準的に進めるための制度である（総務省 2013）。目標管理型政策評価では、省が政策、施策、事務事業からなる政策体系を事前に構築する。施策とは、抽象的かつ長期的な理念である政策と個別具体的な事務事業を結びつける概念である。これまで言及してきた政策庁が企画立案する政策は、目標

管理型政策評価において施策として位置づけられる¹³。政策庁は毎年度の測定および総括的な評価を通じて、施策の有効性を評価する。

まず、目標管理型政策評価では、標準化された様式に基づく測定が年度ごとに行われる。様式とは、事前分析表および評価書を指す。事前分析表は評価書の前に作成される。事前分析表には政策目標、政策手段および測定指標等が記載される。測定指標とは、目標の達成の程度を判断するための指標を表す。その後、政策庁は施策の目標を実現するために実施した自らの活動の評価書の記入を通じて評価する。評価書には、毎年度の活動の実績値、測定指標の値、目標達成度合いの評価、施策の有効性の分析、次年度の施策への評価結果の反映の検討結果等が記載される。

さらに、政策庁は各府省庁が定める政策評価基本計画の期間のなかで一度は施策の総括的な評価を行うように要請されている（総務省 2013）。この評価では、政策庁が施策を実施する前には想定していなかった外部要因の分析、政策手段の有効性および効率性の分析、政策目標の妥当性の判断および以後の政策の企画立案に資する教訓の導出等がされることとなっている。

目標管理型政策評価ではアウトカムが評価される。目標管理型政策評価では、実現されるべきアウトカムが目標として設定され、その実現に向けた進捗状況が毎年度測定される。さらに、基本計画の期間内に一度は、政策目標と政策手段の見直し等を含めた広範な評価が行われる。このように、目標管理型政策評価では、簡便な測定と詳細な調査を組み合わせる形で、アウトカムが測定される。

総合評価とは、施策¹⁴の実施から一定の期間が経過した後に、施策の効果を多角的に調査する方式である。総合評価では政策庁が管理している施策のなかから、政策庁自身の関心に基づいて特定の施策が評価の対象として選択される。そして、施策が社会にもたらしたさまざまな効果が調査される。総合評価は政策庁が任意に行うため、評価の時期は一律に決められていない。

総合評価では、政策庁は広い裁量をもって自らの施策によってもたらされた社会の変化を調

査する。「政策評価の実施に関するガイドライン」(以下、「ガイドライン」という)を見ると、総合評価の対象、実施時点、および評価によって産出すべき情報の内容等の選択に関して、行政組織に広範な裁量が与えられていることがわかる(総務省 2005)。また、評価手法においても、評価対象に応じて定量的または定性的な評価手法を選択すべきであると記載されているが、その選択も行政組織の裁量事項となっている。

総合評価ではアウトカムだけではなく、インパクトも測定可能である。インパクトとは、施策によって社会に生じた波及効果または副次効果といった長期的な変化である。ガイドラインでは、総合評価においては施策の直接的な効果だけではなく、効果が発現するまでの因果関係の連鎖、施策以外で効果の発現に影響を与えた外部要因が調査されるという記載がある。さらに、直接的な効果以外の波及効果または副次的効果までも調査対象になる可能性が指摘されている。

総合評価では政策庁に与えられた裁量が広範であるため、政策庁の機能に適した政策評価の設計も可能になると考えられる。実際に、ガイドラインでは、総合評価は施策の有効性という政策手段の評価だけではなく、政策目標の妥当性も評価される可能性が示されている。また、直接的な効果だけではなく、波及効果または副次的な効果の調査を通じて、政策庁は自らの企画立案が適切であったのか、問題がある場合にはどこに問題があるのかをより深く学習できる。

4. 実施庁と政策庁における評価制度の課題

ここでは、3.で整理した2つの庁の機能に適した政策評価と現行の評価制度との間の違いを参照に、それぞれの庁における政策評価に関する研究課題を提示する。

(1) 実施庁における評価制度の課題

実施庁における政策評価の課題として、第2の評価に関する研究および実績評価における質の確保に関する研究がある。

上述のように、現行の評価制度の下では、実施庁は第2の評価を行っていない。第2の評価とは、実施庁が自らの活動の成果または改善点を把握するための政策評価であった。行革会議が実施庁制度を創設した趣旨は、専門性を備えた実施庁による自律的かつ効率的な政策の実施にあった。こうした趣旨を考慮すると、実施庁は自らの活動の質を向上させるための学習に取り組む必要がある。

そこで、今後の研究課題として、実施庁による第2の評価に関する研究を提示したい。

この研究において、まず第2の評価の要否を判断しなければならない。確かに、実施庁の活動改善に資する情報は、実績評価以外のさまざまな活動によっても産出されうる。たとえば、実績評価以外で独自の評価制度を設計し実施している実施庁がある¹⁵。また、審議会の答申等において、実施庁の活動改善につながる指摘がなされているかもしれない。これらの機会を通じて、実施庁が自らの活動の改善に必要な情報を十分に収集しているならば、第2の評価を実施する意義は低くなる。したがって、実施庁の実施機能の改善に必要な情報を明確にしたうえで、現行の取り組みによって、そうした情報が産出できているのか否かを検討し、実施庁による第2の評価の要否を判断する。

実施庁による第2の評価の実施が望ましいと判断されたならば、その導入に向けた実践的な検討が研究課題になる。この研究では、実施庁による第2の評価における評価目標、評価対象、評価期間、評価手法等といった政策評価の基本設計の構想が目的になる。その際に、実施庁と類似の組織の評価制度を参照するならば、研究成果の妥当性は高まるであろう。参照する組織としては、独立行政法人があげられる。独立行政法人制度は実施庁と同じく、行革会議によって提言された政策の実施を担う組織に関する制度である。独立行政法人が実施するさまざまな評価は、実施庁による第2の評価の構想にあたって有用な例になると考えられる。

また、実績評価における質の確保に関する研究も今後の課題になる¹⁶。政策評価法に基づく政策評価制度および独立行政法人通則法(平成11年

法律第103号)に基づく独立行政法人評価制度は総務省の所管事項であり、総務省はこれらの評価の質の確保に責任を有する。他方で、総務省は各実施庁の実績評価の質を確保するための取り組みを行っていない。総務省が実施した調査によると、この理由は中央省庁等改革の過程で、実施庁ごとに事業の性質が多様であるため、省の長が実施庁の評価の在り方を検討する方が望ましいと判断されたためである(総務省 2004)。そのため、現状の実績評価においては、評価主体に実績評価の運用がゆだねられているが、上述のように、多くの実施庁の評価書は省の長が示した目標、実施庁の活動の実績値、および省の長による判定を簡潔に記載しただけであり、十分な情報が産出されているとはいえない。

こうした問題を研究するにあたって、行政組織が実施する政策評価の質の確保に対して特定の行政組織が責任を有する体制の利点と欠点を明確にする必要がある。続いて、現行の実施庁による実績評価の報告書の質を検討し、質を制度的に確保する必要性を判断する。さらに、統一的な評価制度の制定が望ましいとしても、統一する項目および統一する程度を検討する課題もある。その際に、一部の実施庁によるすぐれた取り組みを統一的に実施することが望ましいのか否かも考慮事項になりうる。

(2) 政策庁における評価制度の課題

政策庁に関しては、その機能である政策の企画立案における学習の必要性を考慮すると、目標管理型政策評価および総合評価が政策庁の学習にどの程度まで貢献しうるのかを明確にする課題がある。

目標管理型政策評価は政府全体的に体系的かつ標準的な実績評価を進めるための評価制度である。上述のように、目標管理型政策評価では府省庁が自らの管理する施策の進捗状況を事前分析表および評価書を用いて、毎年度測定する。これらの標準化された様式は、府省庁による活動量および成果の簡便な把握に貢献する。また、府省庁による施策の評価結果を政策体系に位置づけることで、政府全体の政策、施策、事務事業の進捗状況を内閣、国会、国民が容易に把握

できる。このように、目標管理型政策評価は政府全体的な政策の進捗状況の管理を可能にする取り組みである。

しかし、目標管理型政策評価の目的と政策庁の機能に適した政策評価の目的は異なる。目標管理型政策評価では事前に設定された目標に関連する効果のみに着目する業績測定を手法として用いるため、それ以外の政策の効果に関する評価ができないという指摘がある(北川 2018、p.63)。同様に、山谷も業績測定のみでは政策の内容の改善は生じないと指摘する(山谷 2012、p.223)。これらの指摘を考慮すると、目標管理型政策評価が、個別的な政策領域において不確実性に直面しながらも、高度な創造性を要する政策デザインに取り組む政策庁の能力の向上に貢献する情報を産み出せるのかに関して疑問が生じる。

そこで、目標管理型政策評価を通じた政策庁による学習の可能性が検討課題になる。具体的な方向性として、理論的な考察とすぐれた事例の検討が考えられる。まず、先行研究を参照に、現行の目標管理型政策評価が政策庁の機能の向上に資する情報の産出に貢献するのか否か、貢献するならばどのように貢献するのかを明確にする。つぎに、政策庁が必要とする情報を産出する評価の事例があるのかを調査する。この研究においては、毎年度の測定結果だけではなく、総括的な評価の事例も調査の範囲に含める必要がある。

つぎに、総合評価の活用を通じて、政策庁は自らの機能に適した政策評価の設計が可能になる。上述のように、その理由は総合評価における評価対象および裁量の広さにある。

政策庁による総合評価としては、金融庁に2件、林野庁に1件、観光庁に1件の事例が存在する(金融庁 2005、金融庁 2008、林野庁 2006、観光庁 2011)。総合評価制度が2002年から導入されている事実を考慮すると、政策庁による総合評価の件数は少ない。また、これらの総合評価の目的が政策目標の見直しまたは政策手段の有効性の把握に該当するか否かを確認する必要がある。さらに、それぞれの事例が政策庁の機能に適した政策評価として成功しているのか否かも別に

検討しなければならない。

そこで、総合評価を通じた、政策庁の機能に適した政策評価の可能性の検討が後の研究課題になる。上述のように、制度または理論上は、総合評価を通じて政策庁の機能に適した政策評価が可能である。他方で、その実現のためには、政策庁が自身の政策の特殊性を踏まえて総合評価の方法を個別に設計しなければならない。上述の事例が総合評価の設計に成功しているか否かを検討し、総合評価を適切に設計するためにはどのような事項を検討すべきであるのかを提示する研究を進めることが望ましい。

(3) 本研究以降の展望

本稿の最後に、これまで提示してきた研究課題のなかでも政策庁に適した政策評価の研究が重要であることを主張する。

近年では、内閣が積極的な政策目標を掲げ、その目標の実現に責任を有する組織として政策庁を設置する手法がとられている（鏡 2019）。具体的には、観光庁、スポーツ庁、防衛装備庁および出入国在留管理庁がある。観光庁は2006年の小泉純一郎元内閣総理大臣による観光立国宣言および観光立国推進基本法（平成18年法律第117号）の成立、2007年の観光立国推進基本計画といった自由民主党の政策方針を実現するために設置された政策庁である。スポーツ庁は、官邸主導による2020年のオリンピック・パラリンピックの東京への招致決定後に設置された政策庁である。防衛装備庁は安倍晋三内閣総理大臣のもとでの防衛装備移転三原則の閣議決定を経て、防衛装備の開発と輸出の促進、諸外国との共同研究を担う政策庁として設置された。出入国在留管理庁は、安倍内閣による外国人労働者の受け入れ促進の決定に伴う、出入国管理に関する業務の拡大、外国人の生活の安定化に関する企画立案等の課題に対処する庁として設置された。

このような傾向が続くならば、政策庁の機能と政策評価の適切な関係の検討が、今後の政治および行政の実務においてますます重要になる。これらの政策庁は内閣の政治判断によって設置されているため、政策庁が管理する政策に有効

性があるのかという問いに対して、与党議員、野党議員、関係省庁の職員、研究者、さまざまな利害関係者が関心を抱くであろう。したがって、政策庁の機能に適した政策評価の検討は、実践的に有意義な研究課題になる。以上の理由から、本稿で指摘したいくつかの課題のなかでも、政策庁の機能に適した政策評価の研究が喫緊の課題になると考える。

謝辞

本稿は、日本評価学会春季第16回全国大会における執筆者の報告原稿に加筆修正を加えた論文である。当日の報告では座長の上野宏氏（国際開発センター研究顧問）、多くの方々に有益な御指摘をいただいた。厚くお礼申し上げます。

付記

本稿における意見および主張のすべては、執筆者の個人的見解であり、総務省の見解を何ら示すものではない。

注記

- 1 本稿では政策評価を制度的な観点よりも広い機能的な観点から定義している。具体的には、実施庁の実績評価は行政機関が行う政策の評価に関する法律（平成13年法律第86号。以下「政策評価法」という）における政策評価には該当しないが、本稿の定義には含まれる。
- 2 本稿でいう評価制度とは複数の政策評価に関する制度を指す。具体的には、中央省庁等改革基本法（平成10年法律第103号）に基づく実施庁における実績評価制度および政策評価法に基づく政策庁による政策評価制度が含まれる。
- 3 本稿では用語の簡略化のために、山谷とは異なる名称を概念にあてて。すなわち、「アカウンタビリティ追及」を統制と呼び、「組織の学習」を学習と呼ぶ。

なお、橋本圭多は政策評価の実務における混乱の原因を明確にするために政策評価における監査と評価の考え方を抽出し、検討している（橋本 2017、pp.

- 33-58)。本稿における統制の概念を整理する際に監査の概念を、学習の概念を考える際に評価の概念を参照にした。
- 4 たとえば、障害者政策の性質に着目し、評価の在り方を考察した北川（2018）の研究があげられる。
 - 5 この規定に基づいて、金融庁には局が置かれている。
 - 6 たとえば、国税庁に置かれる国税不服審判所または文化庁における日本芸術院がある。
 - 7 実施庁のなかには、自らの所掌事務に関する企画立案を担う庁がある。公安調査庁、海上保安庁または特許庁があげられる。しかし、実施庁の主な機能である実施の目標は、実施庁自身ではなく所属する省の長が定める点に留意する必要がある。
 - 8 なお、政策庁自身が政策手段を適用するのではなく、所属する省の地方支分部局または地方自治体等に政策手段の適用を委譲する場合もある。この場合には、政策庁が政策手段を適用する過程がなくなる。ただし、政策庁は地方自治体等に権限を委譲したことも含めて自らが企画立案した政策手段が適切であったか否かを評価する必要がある。
 - 9 なお、林野庁のように、政策庁のなかには政策の実施に関する事務を行う庁がある。この事務を行う場合には、政策庁は中央省庁等改革基本法第16条第7項の規定に基づいて、実施庁に準じて、事務の運営の効率化を図らなければならない。
 - 10 一般的に、実施庁には、自らの事業を進めるにあたって必要な範囲で政策対象の行動を強制的に変更させる権限が認められている。政策庁はほとんどの場合、政策手段を通じて政策対象の行動に間接的に働きかける権限しかないため、任務の達成における不確実性が高い。
 - 11 東信男も、実施庁が自らの活動によって政策対象や社会にもたらされた変化を意味するアウトカムを測定すべきであると指摘している（東 2002、pp.206-207）。
 - 12 ただし、国税庁はほかの実施庁による実績評価とはいくつかの点において異なる（国税庁 2019）。たとえば、国税庁は毎年度の評価対象期間を7月1日から6月30日までに設定している。さらに、実績評価の対象になる目標を3段階に区分している。また、評価書において一部、後述する目標管理型政策評価の項目を準用している。しかし、評価の対象が活動の

実績値である点において、ほかの実施庁と共通している。そのため、国税庁の実績評価はほかの実施庁と同様の問題を抱えているといえる。

- 13 以降、これまでの記述の一貫性を確保するために、目標管理型政策評価および後述の総合評価の説明または批判に関する箇所以外では政策庁が企画立案する対象を政策と呼ぶ。また、用語の統一性のために、施策の目標は政策目標、施策を実現するための手段は政策手段と呼ぶ。
- 14 総合評価の対象は、さわめて多様である。「政策評価の実施に関するガイドライン」において記載されているように、総合評価では施策だけではなく、政策から事務事業までの政策体系全体、密接に関連する複数の施策までもが対象になりうる。しかし、本稿では政策庁による1つの施策の改善に資する政策評価を検討するために、総合評価の対象を施策に限定して議論を進める。
- 15 たとえば、山谷（2012b、pp.14-15）は海上保安庁が自ら企画立案した政策の評価を独自で実施している点に注目している。こうした評価が実施機能の改善に資するのかを検討する必要がある。
- 16 三菱UFJリサーチ&コンサルティング（2007）においても同様の課題が指摘されている。

参考文献

- 秋吉貴雄・伊藤修一郎・北山俊哉（2015）『公共政策学の基礎』新版、有斐閣
- 東信男（2002）「政策評価の留意事項と実施状況—ベスト・プラクティスを求めて—」、『会計検査研究』、25:195-218
- 足立幸男（2009）『公共政策学とは何か』、ミネルヴァ書房
- 宇賀克也（2019）『行政法概説Ⅲ—行政組織法／公務員法／公物法—』第5版、有斐閣
- 鏡圭佑（2019）『行政改革と行政責任』、晃洋書房
- 観光庁（2011）『観光立国の推進』、(<http://www.mlit.go.jp/common/000139690.pdf> 2019.9.8最終アクセス)
- 北川雄也（2018）『障害者福祉の政策学』、晃洋書房
- 行政改革会議事務局OB会（1998）『21世紀の日本の行政—内閣機能の強化・中央省庁の再編・行政の減量・効率化—』、行政管理研究センター
- 金融庁（2005）『電子金融取引への金融行政上の対応』、

- (https://www.fsa.go.jp/news/newsj/17/sonota/f-20050831-3/16_s_hyouka.pdf 2019.9.8最終アクセス)
- 金融庁 (2008) 『金融システム改革 (日本版ビッグバン)』、(https://www.fsa.go.jp/seisaku/20_s_hyouka.pdf 2019.9.8 最終アクセス)
- 国税庁 (2019) 『平成 30 事務年度 国税庁実績評価書』、(https://www.mof.go.jp/about_mof/policy_evaluation/nta/fy2018/evaluation/201910ntahyokazentai.pdf 2019.12.15 最終アクセス)
- 総務省 (2004) 『実施庁に係る実績評価に関する調査結果に基づく通知』、(http://www.soumu.go.jp/main_sosiki/hyouka/chousa_040716.html 2019.9.8最終アクセス)
- 総務省 (2005) 『政策評価の実施に関するガイドライン』、(http://www.soumu.go.jp/main_content/000354069.pdf 2019.9.8最終アクセス)
- 総務省 (2013) 『目標管理型の政策評価の実施に関するガイドライン』、(http://www.soumu.go.jp/main_content/000266288.pdf 2019.9.8最終アクセス)
- 中央省庁等改革推進本部 (1999) 『中央省庁等改革の推進に関する方針』、(<https://www.kantei.go.jp/jp/cyuo-syochou/990427honbu/housin.html> 2019.9.8 最終アクセス)
- 橋本圭多 (2017) 『公共部門における評価と統制』、晃洋書房
- 三菱UFJリサーチ&コンサルティング (2007) 『実施庁の実績評価の現状と課題に関する調査』、三菱UFJリサーチ&コンサルティング
- 山谷清志 (2012a) 『政策評価』、ミネルヴァ書房
- 山谷清志 (2012b) 「政策評価の制度化をふまえた理論の再構成—政策評価「学」にむけて—」、『評価クォータリー』、22:11-25
- 林野庁 (2006) 『森林整備目標の進捗状況の検証』、(<http://www.rinya.maff.go.jp/puresu/h18-7gatu/0714b.pdf> 2019.9.8最終アクセス)

(2020.1.15 受理)

Study on the Current Status of and Issues in Policy Evaluation by Agencies

Keisuke Kagami

Office of Administrative Procedure, Administrative Management Bureau,
Ministry of Internal Affairs and Communications.
k.kagami@soumu.go.jp

Abstract

This article presents the issues of the evaluation functions by the agencies which are a form of organization in the Japanese central government. Agencies are classified into two types. First is “implementation agency” that implements project and operational works, and second type is “planning agency” that have a function of the formulation of the program and project. Each agency implements the different evaluations.

In this article, I consider the actual evaluation of each agency. Firstly, I clarify the contents of function of each agency. Based on these functions, I propose the evaluation fitted to each agency's function. Secondly, comparing the suitable evaluation and the actual evaluation, I derive some issue of the actual evaluation.

Keywords

implementation agency, planning agency, performance evaluation, comprehensive evaluation

日本評価学会春季第16回全国大会
「科学的評価と政策の科学」
“Science and Policy in Evaluation”
開催の報告とお礼

2019年5月25日開催（於：JICA地球ひろば）の日本評価学会春季第16回全国大会には、95名の方々にご出席いただきました。誠にありがとうございました。会員各位の日頃の研究や実践活動の報告をもとに、評価研究者、実務者の間の経験、情報、知識を共有化する場として、皆様にとって有益な機会となりましたら幸甚に存じます。多くの方のご尽力により本大会を開催することができましたこと、心より感謝申し上げます。今後とも当学会の活動にご高配賜りますようお願い申し上げます。

実行委員長 正木朋也（国際協力機構）
プログラム委員長 牟田博光（東京工業大学）

2019年5月25日（土）

9:30 - 10:00	受付		
午前の部 10:00 - 12:00	＜共通論題1＞ SDGs評価：未来の世代に、 素晴らしい世界を3 （佐々木亮） Rm600	＜共通論題2＞ 開発途上国支援における 実態評価指標と能力評価指標 （氏川恵次） Rm201AB	＜自由論題1＞ 教育評価 （佐藤由利子） Rm202AB
お昼休み 12:00 - 13:15	お昼休み (12:05 - 13:00 Brown Bag Lunch: APEA大会報告会／廣野良吉・佐藤由利子)		
午後の部Ⅰ 13:15 - 15:15	＜共通論題3＞ IFI（独立財政機関）設立の ためのシナリオ・プランニング・ ワークショップ （石田洋子） Rm600	＜共通論題4＞ JICA評価を政策につなげる 科学的な取り組み：質的・ 量的アプローチの調和 （正木朋也） Rm201AB	＜自由論題2＞ 自治体評価 （田中啓） Rm202AB
午後の部Ⅱ 15:30 - 17:30	＜共通論題5＞ 政策レベルのODA評価 （湊直信） Rm600	＜自由論題3＞ 政策評価・評価制度 （上野宏） Rm202AB	

（ ）は座長です。

日本評価学会春季第 16 回全国大会
「科学的評価と政策の科学」
“Science and Policy in Evaluation”
プログラム詳細

○印は、共同研究の代表者

2019 年 5 月 25 日 (土) 受付 09:30 - 10:00

2019 年 5 月 25 日 (土) 午前の部 10:00 - 12:00

共通論題1 SDGs評価：未来の世に、素晴らしい世界を3				Rm600
		座長	佐々木亮	国際開発センター
K1-1	企業のSDGs貢献に対する評価の現状		家高真衣	国際開発センター
K1-2	SDGs実施・モニタリング評価への技術協力の可能性		石田洋子	広島大学
K1-3	貧困アクションラボ、IPA、3ieのインパクト評価の進捗とSDGsへの貢献		佐々木亮	国際開発センター
共通論題2 開発途上国支援における実態評価指標と能力評価指標				Rm201AB
		座長／モデレーター	氏川恵次	横浜国立大学
K2-1	開発途上国に配慮したSDGsモニタリング手法の開発－廃棄物管理分野の事例－		小島英子	国際協力機構
K2-2	ベトナムにおけるSDGs指標6.3.1モニタリング手法の検討－排水処理分野の事例－		吉田早苗	国際協力機構
K2-3	途上国の廃棄物管理の評価とバランス・スコアカード		加藤郁夫	横浜国立大学
自由論題1 教育評価				Rm202AB
		座長	佐藤由利子	東京工業大学
J1-1	ミャンマー連邦共和国における基礎教育効率化に関する研究		牟田博光	国際開発センター／大妻女子大学
J1-2	JICA青年研修における日本理解への貢献	○	岸本昌子	日本国際協力センター
			菊田怜子	
J1-3	Goal Setting & Case Study of the “Monster” Student’s Change in Professor Soetanto’s Class:-From performance to mastery (目標設定及びケース・スタディ：スタント教室におけるモンスター学生の変化～パフォーマンスからマスタリーへ～)	○	蘇霞	丸和運輸機関
			カワ・スタント	早稲田大学
			ラウ・シン・イー	麗澤大学
J1-4	Reforming Educational Scene in University by Soetanto Method (STM) (スタント法(STM)による大学教育現場の創新)	○	カワ・スタント	早稲田大学
			蘇霞	丸和運輸機関
			ラウ・シン・イー	麗澤大学

2019 年 5 月 25 日 (土) お昼休み 12:00 - 13:15 (12:05 - 13:00 APEA 大会報告会 (Brown Bag Lunch))

Brown Bag Lunch アジア太平洋評価協会 (APEA) 第2回大会報告				Rm202AB
		座長	廣野良吉	成蹊大学／アジア太平洋評価協会
			佐藤由利子	東京工業大学／アジア太平洋評価協会
BBL-1	APEAの活動とAPEA第2回大会の概要		佐藤由利子	東京工業大学／アジア太平洋評価協会
BBL-2	APEA第2回大会開催の意義とGlobal Parliamentarians Forum for Evaluationについて		廣野良吉	成蹊大学／アジア太平洋評価協会
BBL-3	国際協力機構 (JICA) による2つの発表の概要と参加者の反応		河野撰	国際協力機構
BBL-4	APEA第2回大会での発表と参加の感想		宮口貴彰	京都外国語大学

2019 年 5 月 25 日 (土) 午後の部 I 13:15 - 15:15

共通論題3 IFI(独立財政機関)設立のためのシナリオ・プランニング・ワークショップ				Rm600
		座長／モデレーター	石田洋子	広島大学
		副座長／副モデレーター	湊直信	国際通貨研究所
K3-1	分科会「立法府の分析・評価機能の強化に向けて」の背景、目的と活動・ロードマップ		廣野良吉	成蹊大学
K3-2	シナリオ・プランニング・ワークショップの進め方		湊直信	国際通貨研究所
K3-3	IFI設立のためのシナリオ・プランニング・ワークショップ		石田洋子	広島大学
			湊直信	国際通貨研究所
共通論題4 JICA評価を政策につなげる科学的な取り組み：質的・量的アプローチの調和				Rm201AB
		座長	正木朋也	国際協力機構／北里大学
K4-1	ODA事業における質的比較分析(QCA)の応用可能性：JICAインド森林事業の事例	○	氏家慶介	国際協力機構
			角田恵里	CDP Worldwide-Japan
			正木朋也	国際協力機構
K4-2	スリランカ「マムナイ橋梁建設計画」(無償)：「プロセスの分析」による学びの抽出		込山育子	国際協力機構
K4-3	「問題主導型反復実験適用手法」を用いたJICA平和構築事業の実施上の課題解決のための取り組み		田中優子	国際協力機構

○印は共同研究の代表者

自由論題2 自治体評価			Rm202AB	
		座長 田中啓	静岡文化芸術大学	
J2-1	自治体防災政策評価の有効性検証	○	高千穂安長	東北大学大学院(博士後期課程)
			増田聡	東北大学
J2-2	地方自治体文化振興計画に見る政策評価の実態と課題		岩瀬智久	静岡県浜松土木事務所
J2-3	政策評価とAI：理論と制度における可能性を中心に		窪田好男	京都府立大学
J2-4	自治体評価における評価結果の効果的な公表方法		池田葉月	京都府立大学大学院(博士後期課程)
2019年5月25日(土) 午後の部Ⅱ 15:30 - 17:30				
共通論題5 政策レベルのODA評価			Rm600	
		座長 湊直信	国際通貨研究所	
K5-1	外務省ODA評価「外交の視点からの評価」拡充に向けた試行結果		佐々木亮	国際開発センター
K5-2	ODA評価における「外交の視点からの評価」の拡充に向けた取り組みと諸外国の動向		村岡敬一	外務省／OECD/DAC開発評価ネットワーク
K5-3	外務省ODA評価における外交の視点からの評価―「アンゴラ国別評価」を事例に―		桑原恒夫	NTCインターナショナル
自由論題3 政策評価・評価制度			Rm202AB	
		座長 上野宏	国際開発センター	
J3-1	国立大学改革～評価と運営費交付金から捉えた課題～	○	田中弥生	大学改革支援・学位授与機構
			飯間敏弘	東京大学
J3-2	今の日本の大学評価の課題～今の大学と利害関係者のために評価は何をすべきか～		齊藤貴浩	大阪大学
J3-3	庁と評価制度の関係をめぐる予備的考察		鏡圭佑	総務省／関東学院大学
J3-4	アメリカ評価学会 Guiding Principles 改訂の要点		小林信行	OPMAC

日本評価学会春季第16回全国大会セッション報告

共通論題セッション

共通論題1 「SDGs評価：未来の世代に、素晴らしい世界を3」

座長 佐々木 亮（国際開発センター）

2016年から始まった「持続可能な開発目標」(SDGs)は、貧困に終止符を打ち、地球を保護し、すべての人が平和と豊かさを享受できるようにすることを目指す普遍的な行動を呼びかけている。本セッションでは、3つのトピックについて報告して論じ、今後のSDGsの実現について展望した。

冒頭、座長より趣旨説明および共通論題セッションの趣旨と論点の整理を行った。

続いて、「研究報告1：企業のSDGs貢献に対する評価の現状」として、国際開発センターの家高真衣氏から報告があった。その後、「研究報告2：SDGs実施・モニタリング評価への技術協力の可能性」と題して、広島大学の石田洋子氏から報告があった。さらに、「研究報告3：貧困アクションラボ、IPA、3ieのインパクト評価の進捗とSDGsへの貢献」として、国際開発センターの佐々木亮氏から報告があった。

それぞれの報告のあとに、クラリフィケーションの質疑応答があり、3つの発表のあとにフロアからコメントおよび質問をいただいた。それぞれの発表に対して有意義なコメントおよび質問があった。具体的には、「SDGウォッシュにどう対処するのがいいと思っているのか」、「ビジネスセクターにおいても効果的かつ持続的にSDGsに貢献していくためには評価手法の標準化が必要というが、どのように標準化された評価手法があり得るのか」といった質問があった。また、アメリカでは昔から篤志家が設立した財団（Foundations）が大学や研究機関を財政的に支援しており、その伝統が今日のEBPMに資する研究を支えているとの指摘もあった。さらに、現在のSDGsの動きが各国政府主導であることが懸念され、もっと民間部門のアクターを巻き込んでいかなければならないという指摘もあった。

共通論題 2 「開発途上国支援における実態評価指標と能力評価指標」

座長 氏川 恵次（横浜国立大学）

近年の議論として、SDGsのGoal 11およびGoal 12をめぐり、環境・廃棄物管理に関する指標の策定があげられる。こうした問題意識に沿い、今回の報告者2名が所属されている国際協力機構においては、試行的に簡易版としての指標設定を試みるということが一部のプロジェクトで行われている。現在、検討中の段階ではあるが、こうした取り組みについて、本セッションでは具体的な事例研究に基づいた研究報告が行われた。なお、前年度の全国大会からの継承として、SDGs指標の評価についての代表的なツールの1つである、バランス・スコアカードの適用可能性についての検証も、本セッションにおけるもう1つの目的として設定された。

第1報告者である、小島英子氏（国際協力機構）からは、「開発途上国に配慮したSDGsモニタリング手法の開発 - 廃棄物管理分野の事例 -」と題した報告があった。本報告では、開発途上国の諸事情を考慮したSDGsのモニタリング手法のあるべき姿について考察する上で、幾つかの事例としての開発途上国における廃棄物管理の現状と共にタリング上での諸課題、また国際協力機構が関与する専門家会合の論点と、

各手法の開発状況が扱われた。

第2報告者である、吉田早苗氏（国際協力機構）より、「ベトナムにおけるSDGs指標6.3.1モニタリング手法の検討－排水処理分野の事例－」と題した報告がなされた。本報告では、一般に下水処理が普及していない東南アジア、特にベトナムを事例にした調査での結果とインプリケーションが扱われた。具体的には「安全に処理された排水」をいかに定義すべきか、またそのモニタリングのためにはいかなる政策が必要か、といったテーマが対象とされた。

第3報告者である、加藤郁夫氏（横浜国立大学）からは、「途上国の廃棄物管理の評価とバランス・スコアカード」と題した報告があった。本報告は、以下のような仮説の検証を試みた。すなわち「開発途上国支援における実態評価指標と能力評価指標に、バランス・スコアカードを適用することで、カイゼンが見える化でき、長期的で非定量的な情報と短期の定量的な情報を結びつけることが可能となり、社会に貢献するアウトカムを導くことができるのではないか」というものである。

本セッションの位置づけは、2019年12月に、横浜国立大学で開催された、日本評価学会第19回全国大会でのシンポジウム「国内のSDGsに関する活動事例とその評価」を継承・発展するものということである。上記シンポジウムでは、主として国内におけるSDGsの取り組みが議論された。これに対して本セッションでは、開発途上国におけるSDGsの取り組みが対象とされた。今後とも、こうした議論を継承しつつ、その深化を様々な分野に渡って行い、実証を試みていくことが重要であると思われる。

共通論題3 「IFI(独立財政機関)設立のためのシナリオ・プランニング・ワークショップ」

座長／モデレーター 石田 洋子（広島大学）
副座長／副モデレーター 湊 直信（国際通貨研究所）

日本評価学会分科会「立法府の分析・評価機能の強化に向けて」は、2018年1月の設立以来、有識者との勉強会等を通して日本の国・地方自治体の財政赤字を中長期的に削減するための合理的な方法を模索してきた。その結果、国民の代表機関である立法府の積極介入が不可欠で、国会に付置されたIFIの設置が1つの効果的な対策であるという結論に達した。これを受けて、同分科会では、財政赤字に関する日本評価学会をはじめ関係者の考えを理解するためにアンケートを実施すると共に、IFI設立へ向けての準備を開始した。

本セッションは、このIFI設立準備の一環として、日本評価学会春季大会の機会を利用して、分科会の成果を参加者と共有し、本分科会の活動趣旨・進捗への理解を深めると共に、参加型政策立案手法（シナリオ・プランニング）を用いたワークショップにおける参加者との議論を通して、IFI設立へ向けてのより有効なロードマップの作成を目的として実施した。

冒頭、座長より、この共通論題セッションでは、通常の学会発表のセッションとは異なり、上述の通り、シナリオ・プランニング手法による参加型政策立案ワークショップを行うことを説明した。

そして、ワークショップに先立って、成蹊大学の廣野良吉会員から分科会「立法府の分析・評価機能の強化に向けて」の背景、目的、並びに具体的な活動の進捗（勉強会の開催、ロードマップ作成、アンケートの実施と集計等）を報告し、ワークショップのための情報提供を行った。また、国際通貨研究所の湊直信会員が、「シナリオ・プランニング・ワークショップの進め方」と題し、シナリオ・プランニング手法の概要と手順を解説した。

シナリオ・プランニングは10のステップから構成されるが、同セッションにおけるワークショップで全てを行うことはできないので、シナリオ・プランニングの3つ目のステップ「未来を動かす「ドライビング・フォース」を特定する」作業を行った。ワークショップのモデレーターは石田会員、副モデレーター

は湊会員が務めた。

手順としては、まず、IFI設立の背景となっている「政府の増大する債務」を重要課題として、そのドライビング・フォース（原動力、要因、原因）として、歳入面、歳出面、その他の視点から考えられるものについて、同セッションの参加者10名がカードに書いて（1枚のカードに1つの意見）、大判用紙に貼って皆と共有した。参加者から積極的な参加が得られ、合計で100枚近くのカードが貼り出された。モデレーターと副モデレーターが、参加者と議論し、同意を得ながら、重複するカードを整理して、大体の類型化を行った。セッションの最後には、モデレーターおよび副モデレーターが、ワークショップ中の論点を整理し、全ての情報を分科会勉強会で共有する旨を説明した。

同ワークショップで貼り出されたカードを整理した結果、ドライビング・フォースとして以下が挙げられた。

歳入面	労働人口の低下による税収の減少、日本国内への海外からの投資の減少、税制度の問題、景気の低迷、経済成長の停滞、弱体化する産業競争力、企業の生産性向上の停滞、高額所得者の海外への移住、消費の低迷（デフレマインド）、政治家の増税への躊躇、消費税増税の延期、特例国債の発行
歳出面	生活保護世帯の増加、医療費の増加、高齢化、制度設計と現状の人口構成の齟齬、社会保障関係費の増大、予算配分に関する議論の不足、予算審議が適切に行われていない、税金（予算）の無駄遣い、不適切な予算配分、場当たりの補助金、日銀の介入、国債発行の問題、防衛費の増加、途上国への支援、留学生への過剰な保障、自然災害と復興費、社会インフラの老朽化
その他	報道操作、技術力の低下、長期政権、少子化、高齢化、市民の政治への関心の低さ、将来世代のことを考えてない、政治家・官僚の無責任、政治家の財政に関する無知と無関心

限られた時間だったので、十分な検討ができたとは言えないが、参加者10名のうち、3名から分科会に新たに参加したいとの希望が寄せられ、分科会への情報収集を行い、そして分科会の活動を広く知ってもらうという2つの観点から、有益な成果が得られた。ここで得られた情報は、後日開催する分科会の勉強会において共有し、議論を深めていく予定である。

共通論題4 「JICA評価を政策につなげる科学的な取り組み：質的・量的アプローチの調和」

座長 正木 朋也（国際協力機構／北里大学）

国際協力機構（JICA）は2015年以後、毎年秋の全国大会にて共通セッションを企画し、これまでに計4回（述べ5回）の評価関連のセッションを企画開催してきた。これらセッションは、単にJICAの事業評価現場の最新動向を公表するのみならず、参加者との直接的な質疑応答を経て、事業評価システムと評価活動の改善のためのフィードバックをいただくことも目的としている。JICA評価の最新の取り組みのレビューをもとに、今回は、発表者を含め40名ほどの参加者により活発な議論が行われた。

冒頭で座長よりセッション企画の経緯と趣旨と共に、特に今回は質的検討の科学性と政策につながる最新のアプローチを取り上げた背景とその論点を整理した。

第1報告は、氏家慶介氏、角田恵里氏、正木朋也会員（何れもJICA評価部）より「ODA事業における質的比較分析(QCA)の応用可能性: JICAインド森林事業の事例」として、ベン図を用いて質的データを整理し、事業の介入や要因とアウトカムとの因果推論を行うQCAという分析手法のODA事業における適用について報告した。インド森林事業における事業介入とアウトカムとしての森林再生について、QCAの手法

を適用した因果推論の事例を用いてその応用可能性の広さを紹介すると共に、その限界について示した。

第2報告は、込山育子氏（JICA評価部）より、「スリランカ「マンムナイ橋梁建設計画」（無償）：「プロセスの分析」による学びの抽出」として、一昨年から展開している「簡易プロジェクト・エスノグラフィ―手法」の適用事例を報告した。評価の観点から、DAC 5項目では確認できない事実や事業への工夫等の学びの抽出を事業評価に調和させる試みとして紹介した。加えて、2018年10月には「簡易エスノグラフィ―作成ハンドブック」も刊行し、既にJICAホームページで公開していることも共有した。

第3報告は、田中優子氏（JICA評価部）に代わり河野撰氏（JICA評価部）より、「「問題主導型反復実験適用手法」を用いたJICA平和構築事業の実施上の課題解決のための取り組み」として、プロセス分析を考える上で、1つの世界的な潮流ともいえる「問題主導型反復実験適用手法（PDIA）」の適用事例を報告した。効果的なモニタリングとは何か、また紛争影響国等の刻一刻と変化する現場の課題に対し、プロジェクトマネジメントの観点も含め、どのようなアプローチが可能であるかの試みの1つとして紹介した。

参加者からの質問として、第1報告に対しては、介入や要因はどのように選んだか、また回帰分析や相関分析で十分ではないか、等のQCAで用いるデータの背景および分析適用に関わる質疑が集中した。分析目的に応じてどのように対応したかを回答し、定量的なアプローチと異なる視点が必要となることについて補足した。第2報告へは、わざわざプロセス分析にリソースを割く意義は何かとの質問がなされ、演者は運営監理者であるJICA担当者向けの教訓（JICA向けの学び）を深めるためとの回答をした上で、いくつかの限界と共に現在なお試行的に進めている段階であることを補足した。第3報告に対しては、評価のタイミング、様々な評価による情報過多への対処、さらに「発展的評価」においても反復型の問題解決が行われるとして、その関係についての質問がなされた。今回の発表でPDIAは既存のPDMにおけるプロジェクト目標をより具体化するためのアプローチの1つとしての位置づけで紹介した旨の回答がなされた。

全体を通してのコメントとして、「形成的評価（事前設計とは別に事後的にみえてきたものも取り入れ忠実に評価する）」を考える上で、どの演題も非常に興味深いとのコメントがあった。一方で、新たなことに取り組むことそれ自体は良いが、JICA評価システムおよびDAC 5項目評価との関係性を明確にした上で、新たな試みがどの部分の何をどのように改善するかを明確にすべきとのコメントをいただいた。

新たな試みには常に試行錯誤はあるが、評価の全体像を踏まえどのような位置づけやニーズによりそのアプローチを行うかはJICA評価による価値付与の観点から明確にしてゆく必要があること、またそれがJICA評価システムでどのような位置づけにあるかを広く共有できるよう整理して、評価システムの改善と現場ニーズに対応してゆくことが望まれることを確認し閉会した。

共通論題5 「政策レベルのODA評価」

座長 湊 直信（国際通貨研究所）

本セッション「政策レベルのODA評価」では、政策レベルのODA評価において開発の視点と共に実施されている外交の視点からの評価に焦点を当て、外交の視点からの評価についての理解を深め、今後の更なる改善に資することを目的とした。開発の視点からの評価に比べ、その重要性にも関わらず外交の視点からの評価の経験は浅く、知識や経験の蓄積も限られている。本セッションでは、外交の視点からの評価に関して、調査研究の経験、ODA評価ガイドライン等における改善、評価実施者からのコメント等、異なる視点からの報告により理解を深めることを試みた。

第1報告は佐々木亮会員（国際開発センター）による「外務省ODA評価「外交の視点からの評価」拡充に向けた試行結果」である。これは外務省が2018年度に実施した「「外交の視点からの評価」拡充に向けた試行結果」についての報告であり、7つの検討課題が述べられた。すなわち、1. 外交の視点からの評価の目的の明確化、2. 評価対象となる個別のODA政策を通じて貢献が期待される「国益」の内容の確認、

3. 外交と開発の異なる視点から構成される評価の枠組みの維持、4. 一定の標準化された検証方法を用いて事例を積み重ね、5. 分かりやすく内容を伴った結論の記載、6. 具体的な検証に必要な情報入手手段の整備、7. 調査の実施体制の整備、である。

第2報告は村岡敬一会員（外務省／OECD/DAC開発評価ネットワーク）による「ODA評価における「外交の視点からの評価」の拡充に向けた取り組みと諸外国の動向」である。外務省ODA評価において、2010年の「開かれた国益の増進～ODAのあり方検討最終とりまとめ」以降、現在までの「外交の視点からの評価」に関する議論の推移、ODA評価室の取り組み、ODA評価ガイドラインの改善点等について報告が行われた。後半では諸外国の動向として、多くの国々で開発協力が利他主義から国益重視へ移行する傾向とこれを受けた新たな評価の試みが報告された。

第3報告は桑原恒夫氏（NTCインターナショナル）による「外務省評価における外交の視点からの評価－「アンゴラ国別評価」を事例に－」である。外務省が2018年度に実施した「アンゴラ国別評価」を事例に、評価実施者として、外交の視点からの評価に焦点を当てて報告した。外交的な重要性、外交的な波及効果、提言が具体的に解説された。2015年に実施したモロッコ国別評価との比較により、2018年時点での充実された点が明確になった。

全ての報告が終了した後に聴衆との質疑応答が行われた。主な論点は、開発の視点と外交の視点のウェイト付け、他ドナーの外交の視点からの評価の参考例、インタビューにおいて本音をどう引き出せるか、国民への説明責任の視点から外交の視点からの評価の重要性、無償資金協力の評価について、評価手法の研究と外交を専門とする人材による評価等であった。

自由論題セッション

自由論題1 「教育評価」

座長 佐藤 由利子（東京工業大学）

本セッションでは、「ミャンマー連邦共和国における基礎教育効率化に関する研究」、「JICA青年研修における日本理解への貢献」、「スタント法による大学教育現場の創新」、「スタント教室におけるモンスター学生の変化」の4つの貴重な報告があり、活発な質疑応答が行われた。

牟田博光会員（国際開発センター／大妻女子大学）による「ミャンマー連邦共和国における基礎教育効率化に関する研究」では、ミャンマーにおける基礎教育の効率化のため、規模の小さな小学校を、生徒の通学距離等への影響を最小化する形で統廃合する方法をシミュレーションした結果と、それにより予測される成果が報告された。質疑応答では、ミャンマーにおいて村落レベルの小学校が設立された経緯、山や川等の地理的条件の考慮、統廃合による教員一人あたりの生徒数の変化等について質問があった。

岸本昌子会員（日本国際協力センター）と菊田怜子会員による「JICA青年研修における日本理解への貢献」では、2010年～2014年に研修を受けた590名の質問紙調査から、日本・日本人への理解がどのように深まったかを分析した結果が示された。質疑応答では、青年研修参加者の同窓会が、日本の情報を得る重要なチャネルになっていることに関連し、研修参加者の国すべてに同窓会を設立する可能性、同窓会設立による他の効果、国によって同窓会が設立できない理由等について質問が行われた。

カワン・スタント会員（早稲田大学）による「スタント法(STM)による大学教育現場の創新」では、ス

タント会員が考案したスタント法という教育手法の導入で、大学の教育現場がいかに活性化し、学生の授業態度に変化が生まれたかの実例が示された。

また、蘇霞会員（丸和運輸機関）による「目標設定及びケース・スタディ：スタント教室におけるモンスター学生の変化～パフォーマンスからマスタリーへ～」では、スタント会員が早稲田大学で行った英語による授業を受講していた学生の変化について、蘇霞会員がモニタリングした結果が報告された。

2つの報告への質疑応答では、スタント法がどの大学でも適用可能なものなのか、学生が設定した目標を達成できるようにどのような支援が行われたのか、等が質問された。

自由論題2 「自治体評価」

座長 田中 啓（静岡文化芸術大学）

本自由論題セッションにおいては、4組の発表者からそれぞれの研究についての報告があった。報告された内容はいずれも何らかの形で自治体レベルの評価に関連したものであったが、防災、文化政策、AI、評価結果の公表方法等、そのテーマは多岐にわたった。

最初の報告者の高千穂安長会員（東北大学大学院）は、「自治体防災政策評価の有効性検証」という題名の報告を行った（増田聡氏との共著）。同報告では、土砂災害警戒区域を多数抱える広島市を取り上げて、広島市の防災政策の実態を把握し、その妥当性について考察した結果を示した。高千穂氏によれば、広島市のこれまでの防災政策は一定の成果を挙げているものの、経済開発の観点から警戒区域で住宅造営を認めてきたことは、潜在的な被災者を増やしてきた可能性があるという。このことから、防災や経済開発を個別に進めていくだけでは負の効果が大きくなる可能性があるため、さまざまな政策の総和に着目して政策評価を実施する必要性を強調した。

次に、岩瀬智久会員（静岡県浜松土木事務所）からは、「地方自治体文化振興計画に見る政策評価の実態と課題」というタイトルの報告があった。岩瀬会員は、自治体の文化振興計画に示された文化政策の構造的課題を指摘し、さらに、文化政策の評価が行政評価や施設評価に偏っており、政策評価（プログラム評価）が不在となっているとの事実認識を示した。岩瀬氏は、文化政策において政策評価を実施する必要性を述べた上で、地域版アーツカウンシルがその役割を果たす可能性があることを指摘した。地域版アーツカウンシルであれば、文化・芸術や政策評価に関する専門性の高いスタッフにより、継続的に文化政策の政策評価が可能になるという。

第3報告は窪田好男会員（京都府立大学）によるもので、報告の題目は「政策評価とAI：理論と制度における可能性を中心に」であった。窪田氏は近年注目を集めているAIを取り上げ、AIと政策評価の関係について試論的に考察した結果を報告した。まず理論面では、機械学習とディープラーニングを組み合わせることにより、AIが政策の効果等を把握できる可能性があること、言い換えれば、AIが政策評価を行う可能性があるとの見解が示された。一方、制度面では、AIを導入することにより、自治体における評価（自己評価・外部評価）の負担を軽減できる可能性があるとの指摘がなされた。さらに将来的には、自治体が予算管理等に用いる財務管理システムにAIが組み込まれ、系統的に政策評価が実施される可能性についても言及された。

最後に、池田葉月会員（京都府立大学大学院）より「自治体評価における評価結果の効果的な公表方法」という報告があった。池田氏は、自治体における評価結果の公表方法に注目し、現状では評価結果は文章中心の公表方法となっており、そうした公表方法には欠点が多いことを指摘した。またアメリカ合衆国では、インフォグラフィックス、ダッシュボード、マッピング等、視覚的な要素を取り入れて評価結果が公表されている実態についても紹介された。さらに池田氏は、国内の自治体で実施された外部評価の報告書を調査し、その中で視覚的な要素を取り入れている例を把握した結果、評価結果を伝えるための視覚的工

夫は少ないものの、間接的な工夫をしている事例が見られたことも紹介された。ただし、現在の自治体評価の公表方法は、誰をターゲットにしているのかが明確ではなく、相手に応じて効果的に伝えるという点については不十分であるとのことであった。

本セッションでは、各報告が終わるごとにフロアの参加者から質問を受けた。いずれの報告に対しても多くの質疑がなされ、報告者からの応答を含めて参加者間で活発に議論が行われた。

自由論題3 「政策評価・評価制度」

座長 上野 宏（神戸大学国際協力研究科／国際開発センター）

本セッションでは4人の報告者が、評価制度・評価方法・評価者の倫理原則について報告し、フロアの参加者は約40名強であった。

第1報告者の田中弥生会員（大学改革支援・学位授与機構）は、主に国立大学評価制度とそれに基づく国立大学法人への国家予算からの運営費交付金の傾斜配分の現状を記述し、その問題点を明らかにし、今後の方向性について報告した（飯間敏弘氏との共著）。国立大学への評価制度は主には2つある：国立大学法人評価と、重点支援評価である。これら2つの評価制度と運営費交付金に関する問題点から見た、解決の方向性は以下となる：①複数併存し、重複している評価制度の整理・合理化を進めること；②運営費交付金の配分に用いる評価は、明確・共通な判断基準に基づく相対評価が必要である；難しいが、この為の評価方法や指標の開発が求められる；例えば重点支援評価では、2,000以上のKPIの85%がインプットやアウトプット指標であるが、評価指標はアウトカム指標を中心に据えるべきである；③評価者の構成が“身内の評価”となっていると批判されている；大学関係者・専門家ではない第3者民間識者を入れ、国民に説明しうよう構成を見直すこと；④評価結果を運営費交付金へ反映させる場合は、その方法と結果を、最大のステークホルダーである大学関係者へあらかじめ説明すべきである。

フロアからは、今までのような前年踏襲はあり得ないことには賛成するが、具体的にどのように傾斜配分するのか、その配分方法の透明性はあるのか、傾斜配分の結果人件費さえも支払えない地方国立大学が出てくるであろうと言われているが対策はどうなっているのか、等の質問がなされた。日本の将来を左右する重要な制度変更であり、今後もより深い分析・評価がなされることが必要であろう。

第2報告者の齊藤貴浩会員（大阪大学）は主に、国立大学評価制度を経年的にレビューし、それらの問題点を指摘し、それら問題の解決への方向性提言を報告した。対象は、2004年に導入された認証評価（制度）（2004）、国立大学法人評価（2004）、重点支援評価（2016）と、文科省が提言している“共通目標としてのKPIを用いた予算配分のための評価”（2022予定）である。以上のように教育・研究の質向上のための諸評価を行ってきたが、一番の問題点は国立大学の主たる活動である“定常的な教育と研究”が、国立大学法人評価に反映されておらずまた予算配分にも反映されていない点である。第2に、どのように社会や利害関係者（学生、企業）の発展に寄与しているかを社会に示す必要があり、社会に貢献しているというエビデンスを提示する必要があるのではないかと。

フロアからは、評価結果を誰が使うのか？これがはっきりしていないから問題が起きているのではないかと、という質問が出された。答えは、本来は学生、企業、社会が評価結果を利用するはずだが、実際は利用されていない。学生は入学時の偏差値のみに、企業は大学の有名度のみに関心を持ち、社会は大学評価に関心がない。

第3報告者の鏡圭佑会員（総務省）は、諸庁をその機能の観点から実施庁と政策庁に分類し、各類型での機能から考えられるあるべき政策評価制度と現行評価制度との比較検討を報告した。実施庁の現行評価制度は実績評価制度のみで、これは上位の府省が、庁が適正に政策を実施しているかどうかを評価する統

制のための評価である。問題は、実施活動の効果性・効率性を評価するための評価が不足していることである。政策庁の現行評価制度は、目標管理型政策評価と総合評価の2つである。目標管理型政策評価の問題点は、これが府省共通の画一的な評価情報のみを集めるため、庁独自の政策目標の見直しのために有用なのか疑問がある。総合評価の問題は、2002年の総合評価の導入以後今までに総合評価の実施が1件しか存在しないことである。

フロアからは、庁の機能から望ましい評価形態を追及しているが評価利用者からのニーズからも望ましい評価形態が追求できるのではないかと、政策・施策・事業レベルで望ましい評価形態は異なるのではないかと、検討されたのは全て内部評価であるが外部評価を含めると望ましい評価枠組みも変わるのではないかと、との質問がなされた。

第4報告者の小林信行会員（OPMAC）は、アメリカ評価学会（American Evaluation Association）による評価者の倫理規定（guiding principles）につき、既存の2004年版と最新の2018年改訂の比較をし、その主な改訂点の解説とその意味するところの分析を報告した。倫理規定は、①前文と、②基本原則（principles）とそのサブステートメント、から構成されている。最新版では更に、用語集が加えられ、倫理規定中の使用用語（例えば利害関係者）を説明し理解を助ける。前文での変更点は、①冒頭に倫理規定の目的が加えられた；②外部評価向けとの記載が削除され、評価全般に適用されるものとなった；③倫理規定とプログラム評価スタンダードとの関係が規定された；であった。基本原則とそのサブステートメントでは、サブステートメントは各基本原則を説明するものであると位置づけられ、その具体的な行動指針（codes）は大幅に削除された。更に、①体系的調査の意味が明確化された；②作業能力とは、評価者は評価の利害関係者へ、熟達した専門職としてのサービスを提供できること、とされた；③誠実さとは、評価者は、評価の誠実さを確保するために、正直さと透明性を持って行動すること、とされた；④人々への敬意とは、評価者は、諸ステークホルダーの尊厳・幸福・自尊心を尊重し、集団内・集団間における文化の影響を認識すること、とされた；⑤公益と公正とは、評価者は、公平で公正な社会の利益と進展に寄与すべく努力すること、とされた。

フロアからは、common goodの訳は“共通善”ではないのか、原文ではcommon principlesやethical principles等のいろいろな表現が出てきて用語の正式名が判らないがどれなのか等、用語に関する質問がなされた。

Brown Bag Lunch Session

「アジア太平洋評価協会（APEA）第2回大会報告」

座長 廣野 良吉（成蹊大学）
副座長 佐藤 由利子（東京工業大学）

本Brown Bag Lunch Sessionでは、大会の昼食時間を利用し、2019年2月に開催されたアジア太平洋評価協会（APEA）第2回大会の概要、日本評価学会関係者による発表内容、アジア太平洋地域における評価の動向を報告した。

最初に、佐藤由利子会員（東京工業大学／APEA理事）から、「APEAの活動とAPEA第2回大会の概要」

を紹介した。APEAは、2012年9月に設立されたアジア太平洋地域における専門的評価を行う任意団体（Voluntary Organizations for Professional Evaluation, VOPE）のネットワークであり、中心となって活動する団体は、日本評価学会の他、フィリピン、マレーシア、ベトナム、インドネシア、パキスタン、モンゴル、ネパール、スリランカのVOPEであり（APEA第2回大会中、韓国のVOPEの加盟が承認された）、会長は、フィリピン開発評価協会会長のRomeo Santos氏（フィリピン大学教授）である。

APEA第2回大会は、「貧困削減と平和の実現：負託責任、透明性、持続可能な開発のための評価（Reducing Poverty-Enabling Peace: Evaluation for Accountability, Transparency and Sustainable Development）」をテーマとして、2019年2月25日～3月1日にマニラで開催された。初日にはマルコポーロホテルで評価に関するワークショップ、翌26日はフィリピン上院議員会館で、評価のためのグローバル国会議員フォーラム（Global Parliamentarians Forum for Evaluation, GPFE）がアジア、中近東、アフリカの国会議員の参加を得て開催された。27日は、アジア開発銀行で、開会式、基調講演、パネル等が行われ、廣野APEA顧問による基調講演、国際協力機構（JICA）による2つの報告、宮口氏による発表、会場を提供したアジア開発銀行（ADB）中尾総裁によるスピーチ（ADB評価部長による代読）等もこの日に行われた。28日から3月1日にかけては、マルコポーロホテルに会場を移し、基調講演、パネル、発表、ワークショップ等が行われた。期間中に、国際機関や22カ国から202名の参加者があった。

廣野良吉会員（成蹊大学／APEA顧問）からは、APEA第2回大会開催の意義と評価のためのグローバル国会議員フォーラム（GPFE）について、SDGsとの関連で説明が行われた。具体的には、SDGsの実施に当たり、国と国際社会における評価コミュニティへの期待と責任が高まっていること、特に、SDGsの17の目標と国内政策の整合性の確保や、政府と民間のパートナーシップ、中央政府から地方政府／自治体への権限委譲が重要であること、これらに鑑み、APEA加盟国と大会参加国に対し、「国家評価法」を制定し、透明性、負託責任、独立性が確保された評価システムの確立を呼びかけたことが報告された（Hirono, Ryokichi (2019), “The Matters of Peace, Security and Sustainable Development in SDG Evaluation: What Directions are we Heading to?: Self-Questioning on Evaluation Communities in Meeting Effectively to the Growing Threats to People, Peace, Planet and Prosperity in the SDGs Era.”）。

河野摂会員（国際協力機構（JICA））からは、JICAがAPEA第2回大会で行った2つの発表と参加者の反応が報告された。1つ目の発表は、“Results Evaluation and the SDGs: Partnership Modalities for Peace, Security and Sustainable Development” というパネルにおける新井和久評価部長（当時）による「JICAの評価制度」についての発表で、JICAの評価部が他の事業部と並列的に執行部の下に置かれ、評価担当者が事業担当部署との人事ローテーションの中に組み込まれている利点と留意点について紹介した。参加者からは、「現場でも評価スキームを利用しているからこそ、事後評価が生きてくるのだと思う。その点でチームJICAとしての連携を深く感じた」等の感想が寄せられた。

もう1つの発表は、“Organizational Learning: milieus, institutionalization, approaches, and methods in results evaluation” というセッションにおけるJICAフィリピン事務所の海老澤陽次長による「ミンダナオにおけるJICAの平和構築支援」に関する発表で、JICAがミンダナオにおいて実施してきた様々なスキームを「平和と開発」というプログラム目標に沿って整理し、今後のプログラム評価にあたって必要な視点を紹介した。2月22日にミンダナオにおけるパンサモロ暫定政府が発足した直後のタイムリーな発表であったこともあり、会場からは「ミンダナオにおけるJICAのプレゼンスを感じた」等の好意的反応が寄せられた。

宮口貴彰会員（京都外国語大学）からは、APEA第2回大会での発表 “Theory-Based Evaluation Approaches and Critical Realism Concept for Evaluating Complex Interventions（複雑なプログラムの評価のための理論重視型評価手法と批判的实在論の役割と意義）” の概要が報告された。気候変動、防災、環境の持続性、開発（発展）等SDGsを取り囲む複雑な相関関係のあるプログラムの評価には理論重視型の評価手法（TBE）が有用となること、また、という批判的实在論（Critical Realism）が様々なTBEの手法の背景となっていることが紹介された。

委員会活動報告 (2018/2019)

企画委員会

企画委員会は、日本評価学会が行う各種の活動を企画、実行することを通して、我が国の評価学の普及と発展に寄与することを目的としています。最近の活動は以下のとおりです。

1. 顕彰事業

日本評価学会では年1回、学会賞の選定という形で会員の顕彰を行うことにしています。学会賞の受賞者は企画委員会の下に設けられている学会賞審査委員会によって推薦し、理事会で決定されます。学会賞としては、本学会誌に掲載された原著論文等が評価学研究に大きく貢献したと認められる者に授与される「日本評価学会論文賞」、評価学研究の進歩に寄与する優れた研究をなし、将来の発展を期待される者に授与する「日本評価学会奨励賞」、評価学の発展に顕著な功績があった者に授与される「日本評価学会功績賞」の3つの賞があります。

令和元年度は11月1日まで受賞者の推薦を受け付けました。詳細は学会ホーム・ページに記載してあります。受賞者は所定の手続きを経て決定され、12月7、8日に高知大学で開催される第20回全国大会の際に表彰されました。次年度も多くの自薦、他薦を期待しています。

企画委員会委員長 牟田 博光

研修委員会

研修委員会は、我が国における評価学及び評価実践の発展や評価全体のすそ野の拡大に資することを目的に活動を行っています。最近の活動は以下のとおりです。

1. 評価士養成・認定

公共事業、学校教育、福祉、ODA等をはじめとする様々な分野の政策や事業に関するアカウントビリティや効率化が求められており、こうした社会一般からの需要に応えるために、評価システムの導入や事業評価の実施が進められていますが、評価の専門性を持つ人材は未だ十分に育っていないのが現状です。そこで、評価に関する専門的能力を身につけた人材を養成してその能力を認定し、各分野における評価の向上に資することが目的です。

評価士はそのレベルによって初級、中級、上級に区分しています。それぞれのレベルの考え方は次のとおりです。

初級レベル（評価士）： 評価に関する幅広い知識と技法を身につけた専門家

中級レベル（専門評価士）： 各専門分野（セクター）の評価に関する深い知識と技法を身につけた専門家

上級レベル（上級評価士）： 評価に関する体系的な研究を行うことにより、評価活動の発展に寄与することができる専門家

平成30年10月に初級レベルの第25期評価士養成講座を開講し、29人が受講されました。

平成24年度には上級評価士制度を開始し、これまでに17名が認定を受けました。令和元年度は6月20日から8月20日まで募集を行いました。申請はありませんでした。

研修委員会委員長 西出 順郎

出版・広報委員会

出版・広報委員会は、前身の広報委員会に新たに出版機能を加え、2016年12月1日にその活動をスタートさせました。2018年10月以降の主な活動は以下のとおりです。

1. 出版活動

現在、「プログラム評価ハンドブック（仮題）」出版のための編集作業を行っております。執筆者は評価士養成講座の講師が中心となっており、出版後は同講座における活用を予定しております。

2. 広報活動

広報の手段は主として学会ホームページを活用します。従って、学会からのお知らせはホームページに随時掲載します。引き続き、より読みやすく、タイムリーに情報をお伝えするよう努めてまいります。

出版・広報委員会委員長 源 由理子

国際交流委員会

国際交流委員会は、日本が培ってきた評価の経験を活かしながら、評価に関して欧米ならびに途上国との国際的なコミュニケーションを推進することを通じて、国際社会に通用する評価活動の我が国での定着と途上国の評価能力・体制強化に資することを目的としています。最近1年間の活動は以下のとおりです。

1. 第19回全国大会で国際セッションを開催

「評価の国際的潮流」と題して、湊直信（国際通貨研究所）座長のもと、村岡敬一（外務省／OECD/DAC開発評価ネットワーク）がDAC評価項目のSDGsへの適合の議論とその背景、佐藤由利子（東京工業大学／アジア太平洋評価協会）が評価に関する国際的ネットワークとアジア太平洋評価協会（APEA）の概要、石田洋子（広島大学）が北東アジア開発協力フォーラムに参加して：韓国、中国、ロシアのODA評価の現状と課題、をそれぞれ報告した。この3つの報告に対して、聴衆も含めて質疑応答を行った。

2. 春季第16回大会で国際セッションを開催

「政策レベルのODA評価」と題して、湊直信（国際通貨研究所）座長のもと、佐々木亮（国際開発センター）が外務省ODA評価「外交の視点からの評価」拡充に向けた試行結果、村岡敬一がODA評価における「外交の視点からの評価」の拡充に向けた取り組みと諸外国の動向、桑原恒夫（NTCインターナショナル）が外務省ODA評価における外交の視点からの評価「アンゴラ国別評価」を事例に、をそれぞれ報告した。その後、聴衆も参加して、質疑応答、コメントを含む活発な議論が展開された。

3. 第2回APEA conferenceへの参加

第2回APEA大会が2019年2月25日から3月1日までフィリピンで開催された。大会テーマは「貧困削減と平和の実現：説明責任、透明性、持続可能な開発のための評価」で、日本からは、廣野APEA顧問による基調講演、国際協力機構による2つの報告、宮口氏による発表、佐藤理事によるセッション司会、会場を提供したアジア開発銀行（ADB）中尾総裁によるスピーチ（実際はADB評価部長による代読）等が行われ、日本のプレゼンスが見られた大会となった。

4. 春季16回大会でのBrown Bag Lunch Session「APEA第2回大会報告」開催

春季16回大会の昼休みの時間帯に、「APEA第2回大会報告」をBrown Bag Lunch Sessionとして開催した。司会は佐藤由利子（東京工業大学）、発表者は佐藤、廣野良吉（成蹊大学）、河野摂（国際協力機構JICA）、宮口貴（京都外国語大学）の4名で、約80名が参加した。

国際交流委員会委員長 湊 直信

日本評価学会誌刊行規定

2005.2.15改訂

2002.9.18改訂

2001.9.9改訂

(目的および名称)

1. 日本評価学会（以下、「学会」という）は、評価に関する研究および実践的活動の成果を国内外の学界をはじめ評価に関心をもつ個人および機関に広く公表し、評価慣行の向上と普及に資することを目的として、「日本評価研究（仮名）」（英文仮名：“The Japanese Journal of Evaluation Studies”、以下、「評価研究」という）を刊行する。

(編集委員会)

2. 「評価研究」の編集は、後で定める「編集方針」にもとづいて編集委員会が行う。
3. 編集委員会は、学会会員20名以内をもって構成し、委員は学会理事会が選任する。編集委員の任期は2年とし、再任を妨げないものとする。
4. 編集委員会は、互選により委員長1名、副委員長2名および常任編集委員若干名を選出する。
5. 編集委員会は、最低年1回編集委員会を開き、編集方針、編集委員会企画、その他について協議するものとする。
6. 編集委員会は、その活動等について、随時理事会へ報告し、承認を受けるとともに、毎年1回学会年次大会の場で、過去1年の活動成果と翌年の活動計画に関する報告を行う。
7. 委員長、副委員長および常任編集委員は、常任編集委員会を構成し、常時、編集実務に当たる。

(編集方針)

8. 「評価研究」は、原則として、年2回刊行する。
9. 「評価研究」の体裁は、B5版とし、和文又は英文とする。
10. 「評価研究」に掲載する原稿（以下「論文等」という）の分類は、以下の5カテゴリーからなるものとする。
 - (1) 総説
 - (2) 研究論文
 - (3) 研究ノート
 - (4) 実践・調査報告
 - (5) その他
11. 「評価研究」への投稿有資格者は、学会会員および常任編集委員会が投稿を依頼した者とする。学会会員による連名での投稿および学会会員を主筆者とする非会員との連名での投稿は、これを認める。編集委員による投稿はこれを認める。
12. 投稿原稿を上記分類のどのカテゴリーとして扱うかは、投稿者の申請等をもとに常任編集委員会が、下記の「作業指針」に従って決定する。
 - (1) 「総説」は、評価の理論あるいは慣行について概観する論文とし、その掲載については編集委員会が企画・決定する。
 - (2) 「研究論文」は、評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文とし、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
 - (3) 「研究ノート」は、「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考で、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。
 - (4) 「実践・調査報告」は、評価事業の実践あるいは評価にかかわる調査の報告で、その採否については次項に定める査読プロセスを経て常任編集委員会が決定する。

(5)「その他」には、編集委員会が独自に企画する特集に掲載する依頼原稿や学会誌の刊行に関する編集委員会からの学会会員への連絡等が含まれる。

13. 論文等は2名の査読者により査読することとし、その人選は編集委員会が行う。「研究論文」については、査読結果と編集委員会が査読者とは別に指名する担当編集委員1名の参考意見をもとに、編集委員会が掲載に関する決定を行う。「総説」、「研究ノート」、「実践・調査報告」および「その他」の論文については、査読結果にもとづき編集委員会が掲載に関する決定を行う。
14. 編集委員が「評価研究」に投稿した場合には、当該委員はその投稿に係わる常任編集委員会あるいは編集委員会の議事に一切参加しないものとする。
15. 上記いずれのカテゴリーの投稿についても、常任編集委員会による掲載の判断は可・不可の二者択一で行うこととする。但し、場合によっては編集委員会の判断で、小規模の修正による掲載も認める。「研究論文」としての掲載が適当でないと判断された場合でも、投稿者が希望すれば、常任編集委員会は「研究ノート」あるいは「実践・調査報告」としての掲載を決定できる。

(投稿要領の作成公表)

16. 編集委員会は、上記の編集方針にもとづき投稿要領を作成し、理事会の承認を得て、広く公表する。
(配布先)

17. 「評価研究」は、学会会員に無償で配布するほか、非会員に有償で提供する。

(抜刷の配付)

18. 「評価研究」掲載論文等の抜刷り30部を、投稿者（原著者）に無料で配布する。それ以上の部数を希望する場合は投稿者（原著者）の自己負担とする。

(インターネット上の公開)

19. 「評価研究」掲載論文等は、投稿者（原著者）の了承を得て全文をインターネット上で公開する。

(著作権)

20. 「評価研究」に掲載された論文等の著作権は各投稿者（原著者）に帰属するものとし、編集権は本学会に帰属するものとする。

(事務局)

21. 「評価研究」編集及び配布の事務は、それに関連する会計も含めて学会事務局が担当する。

(以上)

『日本評価研究』投稿規定

2019.1.21改訂

2017.7.20改訂

2008.9.29改訂

2003.4.18改訂

2002.3.25改訂

2001.9.9改訂

1. 『日本評価研究』（The Japanese Journal of Evaluation Studies）は、評価に関する論文、論考、調査報告等を掲載する。
2. 『日本評価研究』は、会員間の研究成果交流の場を提供し、内外における評価研究の一層の発展に資することを主目的として発行されており、原則として会員による寄稿を掲載する。なお、依頼原稿を除き、ファーストオーサーは学会員でなければならない。また、投稿は、一時に一原稿に限るとともに、他学会誌などへ二重に投稿などのない未発表のものとする。
3. 投稿された原稿は、編集委員会の責任において審査を行ない、採否を決定する。審査にあたっては、1原稿毎に2名の査読者を選定し、査読結果を参考にする。（査読者には、投稿者名を伏せて査読を依頼する。）
4. 原稿料は支払わない。
5. 『日本評価研究』に掲載された論文等は、その全文をインターネット上の本学会のホームページに掲載する。
6. 投稿にあたっては、原稿の種類が、①研究論文、②総説、③研究ノート、④実践・調査報告、⑤その他のうち、どのカテゴリーに入るかを明記する。ただし、カテゴリーについての最終判断は、編集委員会で行なう。「研究論文」は評価の理論構築あるいは慣行の理解について重要な学問的貢献となると認められる論文、「総説」は、評価の理論あるいは慣行について概観する論文、「研究ノート」は「研究論文」作成過程での理論的あるいは経験的な研究の中間的成果物に相当する論考、「実践・調査報告」は評価事業の実践あるいは評価にかかわる調査の報告、「その他」は編集委員会が独自に企画する特集に掲載する依頼原稿等である。
7. 投稿方法
 - (1) 使用言語は日本語または英語とする。
 - (2) 著者校正は原則として第一校までとする。
 - (3) 英文原稿については、ネイティブスピーカーによる英文チェックを済ませ、完全な英文にして投稿すること。
 - (4) 下記送付先に投稿原稿の電子ファイル（図表を入れた印刷モードのPDFとMS word）をメールで提出する。その際、連絡先（住所、Tel、Fax、Email）と原稿の種類を明記すること。掲載可と判断された原稿については、必要なリライトを経た後に、最終原稿の電子ファイル（MS word）を提出する。その際、オリジナル図表（Excel等）を添付すること。なお、編集に関する最終的

なデザインは編集委員会が行うものとする。

- (5) 日本語原稿の最大文字数は以下のとおり。①研究論文20,000字、②総説15,000字、③研究ノート15,000字、④実践・調査報告20,000字、⑤その他適宜。それぞれ和文要旨を400字程度、英文要旨を150words程度、及び和文・英文でキーワード（5つ以内）を別に添付する。最大文字数20,000字の「研究論文」と「実践調査報告」の場合は、図表を含め刷り上がりが14ページ以内とする。ただし、15～16ページの2ページは著者負担で増頁できる。最大文字数15,000字の「総説」と「研究ノート」の場合は、図表を含め刷り上がりが12ページ以内とする。ただし、13～14ページの2ページは著者負担で増頁できる。左記最大ページ数（16ページと14ページ）を超える投稿は認めない。なお、刷り上がりとは、第1ページの和文要約、第2ページ以降の本文等（図表も含め、20字×43行×2段で設定）、最終ページの英文要約を入れた印刷モードのことである。
- (6) 英文ではA4版用紙に左右マージン30mmをとり、10ポイントフォントを使用し、1ページ43行のレイアウトとする（1ページ約500words）。論文冒頭に150words程度のAbstractをつける。14ページでは、7,000words相当になるが、タイトルヘッド等を考慮して、最大語数を約6,000words（図表、注、文献込み）とする。図表の量によっては、ページ数が予想以上に増える場合もあり得るので、注意すること。ページ数の上限については、日本語原稿の場合と同様である。

8. 送付先

〒108-0075

東京都港区港南1-6-41

品川クリスタルスクエア12階 （一財）国際開発センター内

日本評価学会事務局「日本評価研究」編集担当

E-mail: jes.info@evaluationjp.org

『日本評価研究』執筆要領

2019.1.21改訂

2002.9.18改訂

2002.3.25改訂

1. 本文、図表、注記、参考文献等

- (1) 論文等の記載は次の順序とする。

日本語原稿の場合

第1ページ：表題、著者名、所属先、E-mail、和文要約（400字程度）、和文キーワード（5つ以内）

第2ページ以下（1ページ20字×43行×2段で設定する）：本文（図表含む）、謝辞あるいは付記、注記、参考文献

最終ページ：英文表題、英文著者名、英文所属先、E-mail、英文要約（150words程度）、英文キーワード（5つ以内）

英文原稿の場合

第1ページ：Title; the author's name; Affiliation; E-mail address; Abstract（150 words）; Keywords（5 words）

第2ページ以下：The main text; acknowledgement; notes; references

- (2) 本文の区分は以下のようにする。

例1（日本語）

1.

(1)

①

(2)

(3)

例2（英文）

1.

1.1

1.1.1

1.1.2

- (3) 図表については、出所を明確にする。図表は原則として、筆者提出のものをそのまま写真製版するので、原図を明確に作成すること。写真は図として扱う。

例1：日本語原稿の場合

図1 ○○州における生徒数の推移



(注)

(出所)

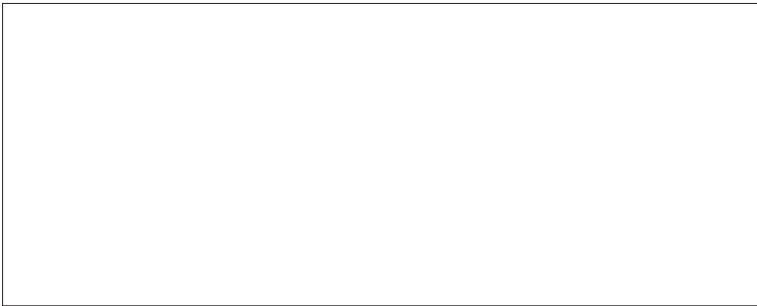
表 1 ○○州における事故件数

--

(注)
(出所)

例 2：英文原稿の場合

Figure 1 Number of Students in the State of ○○



Note:
Source:

Table 1 Number of Accidents in the State of ○○

--

Note:
Source:

(4) 本文における文献引用は、「……である（阿部1995、 p.36）。」あるいは「……である（阿部1995）。」のようにする。英文では、(Abe 1995, p.36) あるいは(Abe 1995)とする。

- (5) 本文における注記の付け方は、(…である¹。)とする。英文の場合は、(…¹)とする。
 (6) 注記、参考文献は論文末に一括掲載する。

注記

1 ……。

2 ……。

- (7) 参考文献は、日本語文献は著者の五十音順、外国語文献は著者のアルファベット順に記し、年代順に記載。参考文献の書き方については以下のようにする。

日本語単行本：著者（発行年）『書名』、発行所

（例）日本太郎（1999）『これからの評価手法』、日本出版社

日本語雑誌論文：著者（発行年）「題名」、『雑誌名』、巻（号）：頁－頁

（例）日本太郎（1999）「評価手法の改善に向けて」、『日本評価研究』、1（2）：3-4

日本語単行本中の論文：著者（発行年）「題名」、編者『書名』、発行所、頁－頁

（例）日本太郎（2002）「行政評価」、日本花子『評価入門』、日本出版社、16-28

複数の著者による日本語文献：著者・著者（発行年）『書名』、発行所

（例）日本太郎・日本花子（2002）『政策評価』、日本出版社

英文単行本：著者（発行年）. 書名. 発行地：発行所.

（例）Rossi, P. H. (1999). *Evaluation: A Systematic Approach* 6th edition. Beverly Hills, Calif: Sage Publications.

英語雑誌論文：著者（発行年）. 題名. 雑誌名, 巻（号）, 頁－頁.

（例）Rossi, P. H. (1999). Measuring social judgements. *American Journal of Evaluation*, 15(2), 35-57.

英語単行本中の論文：著者（発行年）. 題名. In 編者(Eds.), 書名. 発行地：発行所, 頁－頁.

（例）DeMaio, T. J., and Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco, Calif: Jossey-Bass, 177-196.

2名の著者による英語文献：姓, 名, and 姓, 名（発行年）. 書名. 発行地：発行所.

（例）Peters, T., and Waterman, R. (1982). *In Search of Excellence: Lessons from America's Best Run Companies*. New York: Harper & Row.

3名以上の著者による英語文献：姓, 名, 姓, 名, and 姓, 名（発行年）. 書名. 発行地：発行所.

（例）Morley, E., Bryant, S. P., and Hatry, H. P. (2000). *Comparative Performance Measurement*. Washington: Urban Institute.

- （注1）同一著者名、同一発行年が複数ある場合は、(1999a)、(1999b)のようにa,b,cを付加して区別する。

- （注2）2行にわたる場合は2行目移以降を全角1文字（英数3文字）おとしで記述する。

『日本評価研究』 査読要領

日本評価学会 『日本評価研究』 編集委員会

2005年9月10日決定

1. 本査読要領の趣旨

本査読要領は、『日本評価研究』における掲載論文等の審査の要である査読手続きについて、投稿する会員及び査読を依頼される会員に対して解説を行い、審査手続きを効率的かつ効果的に行うことを目的として、定めるものです。

2. 査読の目的と投稿者の責任

査読は、投稿原稿が『日本評価研究』に掲載される論文等としてふさわしいものであるか否かについての判定を当編集委員会が行う上で必要とされるものです。

査読に伴って見いだされた疑問や不明な事項について、必要な場合は修正意見をつけて、修正を求めることがあります。査読は、その意味で、投稿原稿の改善に資するものでもあります。ただし、修正が求められた場合においても、論文等の内容に関する責任は著者が負うべきものであり、査読者の責に帰するものではありません。

査読者は2名で、編集委員会において学会会員の中から当該分野の専門家を選び依頼されますが、学会会員以外に依頼することもあります。

3. 査読の視点

査読は、以下の5つの視点によりますが、投稿原稿の種類によって、重点が違います。

- (1) テーマの重要性・有用度
- (2) 研究の独自性
- (3) 論理の構成
- (4) 実証法・方法論の妥当性
- (5) 評価理論・実践への貢献

- ・研究論文の査読については、上記の5項目全てに配慮する。
- ・研究ノートの査読については、上記5項目のうち、特に (1), (2), (3), (4) の諸項目に配慮する。
- ・実践・調査報告の査読については、上記5項目のうち、特に (1), (3), (5) の諸項目に配慮する。
- ・総説の査読については、上記5項目のうち、特に (3) と (5) の諸項目に配慮する。

4. 投稿に当たっての留意点

2. に掲げた査読の視点以外に、基本的な論文の完成度の問題があります。例えば、

- 論文等として体裁が整っているか、
- 執筆要領にしたがっているか、
- 簡潔明瞭に記述されているか、
- 実証的なデータは適切に位置づけられているか、
- 注や参考文献は本文と対応しているか、
- 専門用語の使用は適切か、
- 語句や文法的な誤りがないか、
- 誤字脱字はないか、
- 句読点に誤りはないか、
- 英文要約などの英文表現は適切か、(必ずしも和文要約の直訳である必要はなく、英文としてまとまっていること)
- 字数は規程に従っているか、

など、内容及び形式に関する留意点があります。

大学院生及び実務家の投稿において、論文としての体裁が整わないまま送付されている例があり、査読そのものに至らないものもあるので、しかるべき指導を受けた後に投稿されるよう強く勧めます。

5. 査読にあたっての判断事例

(1) 完成度において不十分であるが掲載を考慮できる場合

萌芽的な研究、発展が期待できる論文等は評価論の発展のためにできるだけ評価してください。

- 検証は十分とはいえないが、理論や定式化が学問の発展に有用である。
- 考察は十分とはいえないが、新たな理論の形成・促進に有用である。
- 文献調査は十分とはいえないが、研究の位置づけは明確である。
- 比較研究は十分とはいえないが、適用例としては意義がある。
- 考察は十分とはいえないが、社会的、または、歴史的に重要な事例の評価として意義がある。
- 考察は十分とはいえないが、特定の社会活動の評価として意義がある。
- 論文の構成や表現は適切とはいえないが、内容は評価できるものがある。
- 論理性は十分とはいえないが、実務上の有用性がある。
- 有意義な実践・調査報告である。

(2) 掲載を考慮するのが困難と判断される事例

- 問題意識や問題の設定が不明確。
- 基本的な用語の概念の理解や分析枠組が不明確または不適切。
- 論拠とするデータ等の信頼性が乏しい。
- 論旨の明確さや論証の適切さがない。
- 論文の構成、表現（用語、引用、図表等）が適切でない（または整合性がとれていない）。

6. 判定

掲載についての判定は以下の4つの類型に分かれ、最終的に常任編集委員会において決定します。ただし、これらの判定は、評価できる項目や問題のある項目の多少によるものではありません。(3)及び(4)にあるように、投稿論文の種類以外であれば、掲載を考慮できるとする場合があります。別の種類となる場合、字数の関係で、大幅に修正を要することがあります。

(1) 掲載可とする。

(2) 小規模の修正による掲載可とする。

(3) 大幅な修正による掲載可とする。

但し、(総説／研究論文／研究ノート／実践・調査報告)として掲載を考慮できる。

(4) 掲載不可とする。

但し、(総説／研究論文／研究ノート／実践・調査報告)として掲載を考慮できる。

Publication Policy of the Japanese Journal of Evaluation Studies

Last revised on 15th February 2005

The Purpose and the Name

1. The Japan Evaluation Society (hereinafter referred to as “evaluation society”) publishes “The Japan Journal of Evaluation Studies (hereinafter referred to as “evaluation study”) in order to widely release evaluation studies and outputs of practical activities to domestic and international academic societies, interested individual and institutions, and contribute to the advancement and prevalence of evaluation practice.

Editorial Board

2. The editorial board administers editing of evaluation study based on the editorial policy stated below.
3. The editorial board is formed with less than 20 members of the evaluation society who are assigned by the board of directors. Terms of editors are two years but can be extended.
4. The editorial board assigns one editor-in-chief, two vice-editors-in-chief, and a certain number of standing editors among the members.
5. The editorial board may hold at least one meeting to discuss the editing policy, plans of editorial board, and others.
6. The editorial board reports activities to the board of directors as needed and receives approval. Also it is required to report the progress of the past year and an activity plan for the following year at the annual conference.
7. The editor-in-chief, the vice-editors-in-chief and the standing editors organize the standing committee and administer editing on a regular basis.

Editorial Policy

8. The evaluation study, as a principle, is published twice a year.
9. The evaluation study is printed on B5 paper, and either in Japanese or English.
10. Papers published in the evaluation study are categorized as five types;
 - 10.1. Review
 - 10.2. Article
 - 10.3. Research note
 - 10.4. Report
 - 10.5. Others
11. The qualified contributors are members of the evaluation society (hereinafter referred to as “members”) and persons whose contribution is requested by the standing editors. Joint submission of members and joint submission of non-members with a member as the first author are accepted. Submission by the editors is accepted.
12. Submitted manuscripts are treated as the above categories, however, the standing editors will decide based on the application of the contributors and the following guidelines;
 - 12.1. “Review” is a paper, which provides an overview of evaluation theory or practice. The editorial board will make the decision regarding publication.
 - 12.2. “Article” is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.
 - 12.3. “Research note” is a discussion equivalent to the intermediate outputs of a theoretical or empirical enquiry. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.

- 12.4. "Report" is the study report related to a practical evaluation project or evaluation. The standing editors committee makes adoption judgments following the referee-reading process described in the next section.
- 12.5. "Others" includes requested papers for special editions organized by the editorial board and announcements from editorial board to members regarding publication.
13. The editorial board selects two referee readers. For the "article", the editorial board makes adoption judgments referring to the results from referee readings and comments provided by one editor assigned by the editorial board. For "review", "research note", "report" and "others", the editorial board makes adoption judgments referring to the results from referee readings.
14. When editors submit a manuscript, the editors are not allowed to attend any of the standing editors committee meetings or editorial board meetings regarding the manuscript.
15. The standing editors have alternative of approval or not-approval for adoption judgment of manuscripts submitted to any categories. However exception is permitted if the editorial board approves the publication after minor rewrite. Even if the manuscripts are considered insufficient as an "article", standing editors can decide whether the manuscripts are published as a "research note" or "report" if the authors wish to publish.

Formulation and Release of Submission Procedure

16. The editorial board formulates the submission procedure based on the editorial policy described above and release after approval from the board of directors.

Distribution

17. The evaluation study is distributed to all members for free and distributed to non-members for a charge.

Distribution of the Printed Manuscript

18. 30 copies of the respective paper are reprinted and distributed to the authors. The authors must cover any costs incurred by author's requests for printing more than 30 copies.

Release on the Internet

19. The papers published in the evaluation study are released on the internet with approval from the authors.

Copyright

20. Copyright of papers which appear in the evaluation study is attributed to the respective authors. Editorial right is attributed to the evaluation society.

Office

21. The office is in charge of administrative works for editing, distribution, and accounting.

Information for Contributors (For English Papers)

Last revised on 20th July 2017

1. "The Japanese Journal of Evaluation Studies" is the publication for reviews, articles, research notes, and reports relating to evaluation.
2. "The Japanese Journal of Evaluation Studies" is primarily published to provide opportunities for members of the Japan Evaluation Society (hereinafter referred to as "members") to exchange findings, and to contribute to further development of the study of evaluation both domestically and internationally. As a principle, this journal publishes the contributions submitted by the members. With the exception of requested papers, the first author must be a member. A submission (as the first author) is limited to one manuscript that has not been published or submitted in any form for another journal of academic association etcetera.
3. Adoption judgments of the manuscript are made at the discretion of the editorial board. Comments from two referee readers who are appointed for every manuscript are referred to in the screening process (the editorial board requests referee readers without notifying the author of manuscript).
4. Payment for the manuscript is not provided.
5. Papers published in "The Japanese Journal of Evaluation Studies" are released on the Internet at homepage of this academic society.
6. Regarding submission, manuscripts must be identified as one of the following categories: 1) article, 2) review, 3) research note, 4) report, and 5) others. However, the final decision of the category is made by the editorial board.
 "Article" is considered as a significant academic contribution to the theoretical development of evaluation or understanding of evaluation practice.
 "Review" is a paper which provides an overview of evaluation theory or practice.
 "Research note" is a discussion equivalent to the intermediate outputs of a theoretical or empirical study in the process of producing an "article".
 "Report" is the study report related to a practical evaluation project or evaluation.
 "Others" are manuscripts for special editions requested by the editing committee.
7. Manuscript Submission
 - (1) Manuscripts may be written in either Japanese or English.
 - (2) Correction by the author is only for the first correction.
 - (3) English manuscripts should be submitted only after the English has been checked by a native speaker.
 - (4) Submit manuscripts via email. Contact information including mailing address, telephone number, fax number, e-mail address, and the category of the manuscript should be clearly stated.
 For approved manuscripts, after necessary rewriting, the author needs to submit the final paper via email. Original figures, charts, and maps should be provided.
 - (5) Total printed pages should not exceed 14 pages. Any cost incurred by printing more than 14 pages must be covered by the author.
 - (6) The layout for English papers should be 30 mm of margin at left and right side, 10pt for font size, 43 lines on A4 paper (about 500 words per page). An abstract of 150 words should be attached to the

front. 14 pages are equivalent to 7,000 words but the body should not exceed 6,000 words to allow for the title, header, figure, chart, footnotes, and references. Please note that the number of pages may be more than expected depending on the number of figures included.

8. Mailing address

Office of Japan Evaluation Society at International Development Center of Japan
Shinagawa Crystal Square 12th Floor, 1-6-41 Konan, Minato-ku, Tokyo,
108-0075, Japan
E-mail: jes.info@evaluationjp.org

Writing Manual of the Japanese Journal of Evaluation Studies (For English Papers)

Revised on 18th September 2002

1. Text, Charts, Figures, Graphs, Diagrams, Notes, and References

(1) The paper should be written in the follow order:

First page: Title; the author,s name; Affiliation; E-mail address; Abstract (150 words); Keywords (5 words)

Second page: The main text; acknowledgement; notes; references

(2) Section of the text should be as follow:

1.

1.1

1.1.1

1.1.2

(3) Source of the charts, figures, graphs, and diagrams should be clarified. Submitted charts and others will be pzhotoengraved, therefore it is important that the original chart is clear. Pictures shall be treated as figures.

Figure 1 Number of Students in the State of ○○



Note:

Source:

Table 1 Number of Accidents in the State of ○○

--

Note:

Source:

(4) Citation of literature in the text should be, (Abe 1995, p.36) or (Abe 1995).

(5) Note in the text should be, (-----.¹)

- (6) Note and references should be written all together in the end.

Note

1 -----.

2 -----.

- (7) Reference should list the literature in alphabet order, and arranged in chronological order. Follow the examples:

Book: author (year of publication). *Title of the book*. Published location: publishing house.

(e.g.) Rossi, P. H. (1999). *Evaluation: A Systematic Approach 6th edition*. Beverly Hills, Calif: Sage Publication.

Article from magazine: author (year of publication). Title. *Title of the magazine*, volume (number), page-page.

(e.g.) Rossi, P. H. (1999). Measuring social judgments. *American Journal of Evaluation*, 15(2), 35-37.

Article in Book: author (year of publication). Title. In editor (Eds.), *Title of the book*. Published location: publishing house, page-page.

(e.g.) DeMaio, T. J., and Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Fransisco, Calif: Jossey-Bass, 177-196.

Book by two authors: surname, first name, and surname, first name. (year of publication). *Title of the book*. Published location: publishing house.

(e.g.) Peters, T., and Waterman, R. (1982). *In Search of Excellence: Lessons from America,s Best Run Companies*. New York: Harper & Row.

Book by more than three authors: surname, first name, surname, first name, and surname, first name. (year of publication). *Title of the book*. Published location: publishing house.

(e.g.) Morley, E., Bryant, S. P., and Hatry, H. P. (2000). *Comparative Performance Measurement*. Washignton: Urban Institute.

(note 1) If some references are from the same author with the same publication year, differentiate by adding a,b,c as (1999a), (1999b).

(note 2) If the reference is more than a single line, each line from the second should be indented by three spaces.

(e.g.) DeMaio, T. J., and Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Fransisco, Calif: Jossey-Bass, 177-196.

Referee-Reading Guideline

The Japanese Journal of Evaluation Studies Editorial Board,

The Japan Evaluation Society

Approved on 10th September 2005

1. Content of the Referee-Reading Guideline

This Referee-Reading Guideline is to provide explanation of the main publication judgment, procedure of the referee-reading, to the members who submit the manuscript and for the members who are requested to conduct referee-reading in order to carry out the procedure efficiently and effectively.

2. Purpose of Referee-Reading and the Responsibility of the Author

Referee-reading is necessary for the editorial board to make decisions of whether submitted manuscripts are appropriate to publish in the Japanese Journal of Evaluation Studies or not.

If there is doubt or obscurity identified in manuscripts during the referee-reading corrections may be required. Therefore, referee-reading also contributes to the improvement of the submitted manuscripts. However, although the manuscripts are requested corrections, the author is still solely responsible in regards to the contents and it is not attributed to the referee-readers.

Referee-readers are two persons who are requested by the editorial board depending on the specialty or the field of the submitted manuscript. People who are not members of this academic society also may be requested.

3. Items of Consideration in Referee-Reading

Five points are considered in referee-reading, however, the importance of each may be different depending on the type of manuscript.

- (1) Importance and utility of the theme
- (2) Originality of the study
- (3) Structure of the logic
- (4) Validity of verification and methodology
- (5) Contribution to evaluation theory and practice

- For the article, all of above five are considered.
- For the research note, especially (1), (2), (3), and (4) are considered.
- For the report, especially (1), (3), and (5) are considered.
- For the review, especially (3) and (5) are considered.

4. Attentions in submission of manuscript

Besides above five viewpoints, basic completeness as a paper is also considered, for example;

- appearance of the paper is organized
- written according to the writing manual
- described simply and distinctive
- verification data is appropriately used
- notes and references are corresponding with the text
- terminology is appropriately used
- no wording and grammatical mistakes
- no errors and omission
- no punctuation mistakes
- expression in English abstract is appropriate
- word count is according to the manual

The above mentioned forms and contents are also considered. There have been cases in which graduate students and practitioners posted without organizing the manuscripts as a paper. On those occasions, referee-reading was not conducted. Necessary consultation is strongly recommended prior to submission.

5. Judgment Cases in Referee-Reading

(1) In the case of the manuscript which is considered acceptable for the publication but is not yet complete:

The referee reader should evaluate carefully whether the paper can contribute to the development of evaluation theory or evaluation studies.

- Verification is lacking but the theory and formulation are useful for academic development.
- Analysis lacking but useful for formation and promotion of new theory.
- The literature review is not of a high standard but, the overall study is meaningful.
- Comparative study is not up to standard but is meaningful as an example of application.
- Analysis is lacking but it is meaningful as an evaluation of socially and historically important cases.
- Analysis is lacking but it is meaningful as an evaluation of particular social activities.
- Organization and expression are not up to standard as a paper but the contents are worthy to evaluate.
- Logic is not strong enough but useful in practice.
- The paper has significance as a report.

(2) In case of the manuscript which is considered as difficult for publication:

- Awareness of the issue or setting of the problem is indecisive.
- Understanding or analytical framework of notion of basic terminology is indecisive or inappropriate.
- There is a lack in credibility of data for the grounds of an argument.
- There is no clear point of an argument or appropriateness of proof.
- Organization of the paper and presentation (terminology, citation, chart, etc) are inappropriate (or not consistent).

6. Judgment

The final decision will be made on publication at the standing editors committee following one of four patterns (listed below). However, these judgments are not based on the number of errors but on the strength of the overall report. In the case of (3) and (4), there is a possibility to be published as a different type of paper. If it is published as a different type of paper, major rewrite concerning the number of words may be required.

- (1) The paper will be published.
- (2) The paper will be published with minor rewrite.
- (3) The paper will be published with major rewrite, however as a different type of paper (review, article, research note, or report).
- (4) The paper will not be published; however there is the possibility that it will be published as a different type of paper (review, article, research note, or report).

～投稿案内～

日本評価学会では、「日本評価研究」掲載のための投稿原稿を募集しております。随時、投稿を受け付けておりますので、ご興味をお持ちの方は投稿規定・執筆要領をご参照のうえ、奮ってご投稿ください。<http://evaluationjp.org/activity/journal.html>

投稿先：特定非営利活動法人日本評価学会編集委員会

〒108-0075 東京都港区港南1-6-41 品川クリスタルスクエア12階

一般財団法人国際開発センター内

E-mail : journal@evaluationjp.org

『日本評価研究』第20巻第1号

2020年3月31日

編集・発行

特定非営利活動法人日本評価学会編集委員会

〒108-0075 東京都港区港南1-6-41

品川クリスタルスクエア12階

一般財団法人国際開発センター内

印

刷

株式会社 研恒社

©日本評価学会

本誌に掲載されたすべての内容は、日本評価学会の許可なく転載・複写はできません。

Japanese Journal of Evaluation Studies

Vol. 20, No. 1, March 2020

CONTENTS

Research Note

Considerations on Respect for Evaluation Participants in Evaluation Ethics:
Implications from Research Ethics in Psychology Nobuyuki Kobayashi

Study on the Current Status of and Issues in Policy Evaluation by Agencies
Keisuke Kagami

Report of the 16th Spring Conference

Report of the Committees