

日本評価研究

Japanese Journal of Evaluation Studies

Vol. 21, No. 2, September 2021

巻頭言

評価学に関心をもつ多くの関係者が集い、活発に交流できる学会運営を
目指して 大島 巖

特集：持続可能開発目標の評価(SDGs)

特集に寄せて「持続可能開発目標の評価(SDGs)」 林 薫

Evaluation Capacities to Advance Sustainable Development for All
Indran A. Naidoo

SDGsの時代における評価能力構築 村岡 敬一 日野 類子

SDG教育目標にみる理念志向ターゲットの評価に関する一考察：
測定可能性(measurability)から評価可能性(evaluability)へ
米原 あき

SDGs第7目標の評価 林 薫

DAC新評価基準の解釈と運用 江口 雅之

新規SDG事業構想時の評価活用：
アフリカの次世代企業家育成事業の事例報告 長尾 真文

研究ノート

プログラム評価の一類型としての「社会的インパクト評価」の課題と可能性
伊藤 健 玉村 雅敏 植野 準太

留学生30万人計画の成果と課題
－成長戦略、大学のグローバル化及び日本語教育との関係からの考察－
佐藤 由利子

実践的判断のプロセスとしての形成的アセスメント
－J.デューイの価値評価論をめぐって－ 西塚 孝平

第22回全国大会のご案内

『日本評価研究』編集委員会
Editorial Board

編集委員長
Editor-in-chief

西野 桂子(関西学院大学)
Keiko NISHINO

常任編集委員
Standing Editors

大島 巖(日本社会事業大学)
Iwao OSHIMA

牟田 博光(東京工業大学)
Hiromitsu MUTA

山谷 清志(同志社大学)
Kiyoshi YAMAYA

米原 あき(東洋大学)
Aki YONEHARA

編集委員
Editors

窪田 好男(京都府立大学)
Yoshio KUBOTA

贅川 信幸(日本社会事業大学)
Nobuyuki NIEKAWA

橋本 昭彦(国立教育政策研究所)
Akihiko HASHIMOTO

林 薫(文教大学)
Kaoru HAYASHI

村上 裕一(北海道大学)
Yuichi MURAKAMI

特定非営利活動法人日本評価学会
編集委員会事務局
E-mail: journal@evaluationjp.org

日本評価研究

第21巻 第2号 2021年9月

目次

巻頭言

大島 巖

評価学に関心をもつ多くの関係者が集い、活発に交流できる学会運営を目指して…………… 1

特集：持続可能開発目標の評価（SDGs）

林 薫

特集に寄せて「持続可能開発目標の評価（SDGs）」…………… 3

Indran A. Naidoo

Evaluation Capacities to Advance Sustainable Development for All…………… 5

村岡 敬一 日野 類子

SDGsの時代における評価能力構築…………… 19

米原 あき

SDG教育目標にみる理念志向ターゲットの評価に関する一考察：

測定可能性（measurability）から評価可能性（evaluability）へ…………… 31

林 薫

SDGs第7目標の評価…………… 47

江口 雅之

DAC新評価基準の解釈と運用…………… 63

長尾 眞文

新規SDG事業構想時の評価活用：

アフリカの次世代企業家育成事業の事例報告…………… 73

研究ノート

伊藤 健 玉村 雅敏 植野 準太

プログラム評価の一類型としての「社会的インパクト評価」の課題と可能性…………… 89

佐藤 由利子

留学生30万人計画の成果と課題

ー成長戦略、大学のグローバル化及び日本語教育との関係からの考察ー…………… 103

西塚 孝平

実践的判断のプロセスとしての形成的アセスメント

－J.デューイの価値評価論をめぐって－ 117

第22回全国大会のご案内 131

【巻頭言】

評価学に関心をもつ多くの関係者が集い、 活発に交流できる学会運営を目ざして

大島 巖

日本評価学会会長／東北福祉大学副学長・教授

昨年末に行われた日本評価学会理事会・総会において期せずして会長にご選任頂き、早いもので本誌発刊時には9ヶ月が過ぎようとしています。20年以上の歴史を持つ重要な学会の会長という重責を担うこととなり、この間身の引き締まる思いで学会運営に取り組んで来ました。

私の学問的なバックグラウンドは社会福祉学・保健学であり、評価学領域では自分の専門に引き寄せながら、実践家参画型の形式的評価、エビデンスに基づく実践（EBP）などに取り組んで来ましたが、本学会の中では未だマイナーな存在と認識しています。そのような私を会長に選任して頂いた、学会組織から期待される私の重要な「使命」は、昨年以来、前会長の山谷清志先生や他の役員の皆さんとご一緒に取り組んで来た学会改革を、今期役員任期のうちにひとまずの区切りを付けること、そしてそれを次の執行部に発展的に引き継ぐことにあると認識しています。

ご承知のように、昨年4月には学会事務局を、20年間お世話になった国際開発センターから行政管理研究センターに移転しました。この移転を契機に、前執行部は事務局機能の大幅な見直しを行い、理事活動を中心とした学会運営へと大きくシフトを切り替えて来ました。私の任期中には、この学会運営の改革を組織として定着させて、より魅力ある学会運営が行えるよう、次の執行部に対して方向付けをし、バトンタッチしていきたいと考えています。

この改革によって、学会運営を会員の皆さまの主体的参加による、より活発で自律的なものにすること、また開かれた学会運営を行うことにより、会員の皆さまのさまざまなご経験を活かして、相互交流が進展することを目ざしたいと思っています。

お陰さまで、現執行体制が発足した昨年末以来、新役員の皆さまにはたいへん精力的かつ献身的にご活動頂き、将来の方向性を一定程度展望できるところまで辿り着いたように考えております。この場を借りて、関係された皆さまには心より感謝とお礼を申し上げたいと思います。

新体制での理事会は、これまで計3回（2020年12月、2021年3月・6月）開催しました。また正副会長と6つの常設委員会の委員長から構成される常任理事会は、これまでに計5回（2020年12月、2021年1月・2月・4月・6月）開催しています。常任理事会開催の間には、正副会長と事務局による運営・調整ミーティングと、各委員会が複数回の会議を持ち、常任理事会・理事会では、学会全体の調整と合意を得る方式で、学会運営を行って来ました。

委員会活動をこれまで以上に開かれた、活発なものにすることが、これまでの本学会の大きな課題でした。各委員会には、原則として他の委員会との併任のない専任担当理事を複数名おき、加えて理事以外の委員会委員にはそれぞれの領域で活躍されている中堅・若手の会員を各委員長からご挙挙頂いて関わって頂くこととしました。このことが、委員会活動の活性化につながって来たと考えております。

さらには、各委員会には評価学を構成する多様な領域を代表する会員をバランス良く配置することを心

がけました。評価学は学際的な学問です。本学会では政策評価、開発評価の領域から多くの人材が現在会員になっています。その他の領域は会員数が必ずしも多くありませんが、保健・医療・福祉・教育・労働・刑事司法など対人サービス領域における評価は世界的に見るととても重要です。本学会においても、この「第3の領域」についても、ある程度は意識的に委員を配置しました。

以上の活動の中で、本学会の主要3事業である、①学会大会の開催（企画委員会所掌）、②学会誌の発行（編集委員会所掌）、③評価士養成講座の開催（研修委員会所掌）を含めて、学会運営の多くの活動を、6つの常設委員会が中心になって担えるようになったと考えています。

さて、以上の委員会活動に加えて、3名の副会長には学会運営改革の重要課題について、会長特命事項の検討をお願いしました。

新体制が発足して副会長に付託した特命事項は、5項目ありました（副会長担当は3月理事会時のもの）。
 《1》会員ニーズに応える学会財政の見直し：年会費の見直し、会員増への戦略作成、経費削減（学会誌発行に関わる経費削減、大会経費の削減など）、契約業務のルール化など（石田副会長担当）

《2》役員選挙制度のあり方（石田副会長担当）

《3》評価士の体系的な資格制度のあり方（津富副会長担当）

《4》会員の倫理と権利に関する規程整備（研究倫理規程のあり方、著作権関係など）（津富副会長担当）

《5》春季大会の持ち方、分科会の活性化（齊藤副会長担当）

各副会長には、精力的に検討を進めて頂き、本年6月に開催した常任理事会・理事会には、中間報告を含めた答申レポートをご提出頂きました。

このうち、石田副会長に諮問した《1》項目と、齊藤副会長に諮問した《5》項目への答申報告は、今後の本学会活動基盤に関わる重要な内容を多く含んでいます。具体的には、(1) 会員数拡大の方策、(2) 会費値下げの可能性、(3) 学会誌の発刊形態（電子ジャーナル化等）や刊行コスト削減、(4) 大会の実施形態と参加費徴収方法、(5) 会員サービス・情報発信の拡充、(6) 研修事業の拡大可能性、(7) 委託事業の見直しなどです。

この中間答申を踏まえて、6月理事会では、常任理事会の中に「学会の新たな発展に向けた基盤整備検討会」を設置することとしました。

この検討会では、今後、

- ・2021年12月までに活動方針中間報告のまとめ
 - ・2022年9月までに活動方針最終報告のまとめ
- を、ご報告頂くことを予定しています。

以上を踏まえて、いまの役員任期のうちに、本学会の今後の新たな発展に向けた基盤整備の方向性を、来年の理事会・総会に対して提示して行きたいと考えております。

日本の評価学をめぐる近年の新しい情勢として、既存制度では対応できない様々な新しい社会課題の拡大に対して、課題解決のために新たな発想で生み出される「変革的プログラム」への評価に対する関心が高まっています。2016年に制定された休眠預金等活用法では、評価が制度の中に明確に位置づけられました。また近年、社会的インパクト評価やエビデンスに基づく政策立案（EBPM）に関心が高まり、科学的な方法で「変革的プログラム」をより効果的なプログラムモデルへ発展させる有効な評価法が求められています。

このように評価学に対する社会の期待が高まる中、本学会の会員数のさらなる拡大や、会員ニーズにより十分に応え得る学会活動の展開が望まれます。

評価学に関心をもつ多くの関係者が本学会に集い、活発に交流できる学会運営を行うことによって、意欲があり意識の高い人材が、より多く本学会にご参画頂くことで、本学会をさらに発展させて行きたいと考えております。

会員の皆さまの引き続きのご理解・ご協力と学会活動への積極的なご参画を宜しくお願い申し上げます。

【特集に寄せて】

持続可能開発目標の評価（SDGs）

林 薫

文教大学国際学部

持続可能な開発目標（Sustainable development Goals以下SDGsと略）が2016年にスタートしてから早くも5年が経過し、15年間の「中盤戦」に突入している。あと2年もすれば中間年に到達する。早晩、SDGsの達成見込みや2030年以降の目標の議論が開始されなければならない時期にきている。SDGsの目標は達成しうるのか、SDGsに続く目標はどのようなものになるべきかについて早晩世界的な議論が始まることになる。

SDGsのこれまでの歩みは平坦ではない。開始直後に英国のEU離脱、米国のトランプ政権の成立などに代表されるような一国主義の流れが顕著になり、SDGsがその実施の前提とする国際協力・国際協調体制に逆風が吹きはじめた。ロシア、中国、ミャンマー、ベラルーシのように、SDGs第16目標の「安定して包摂的な法制度と法の支配」という目標に真っ向から反する政治的事件が頻発している。SDGsとほぼ同時に成立し、はじめて主要国が参加して成立したパリ協定からトランプ政権が一時離脱を表明したことも暗い影を投げかけた。加えて、2020年初頭から始まったコロナ（COVID-19）パンデミックは、全世界に例外なく襲い掛かり、多くの命を奪うとともに、経済を立ちいかなくさせ、途上国の貧困層に特に深刻な打撃を与えている。SDGs第1目標の貧困解消目標は勿論のこと、コロナ対策によってリソースがクラウド・アウトされる結果として、他のすべての目標の実現のためのリソースが不足し、目標の達成が危うくなっている。

一方で、好ましいと思われる展開がないわけではない。まず、米国バイデン政権はパリ協定に復帰し、国際協調の機運も回復の兆しがある。コロナ禍のもと「全員が安全にならないうちは誰も安全ではない」という認識が広まり、ワクチン提供の国際的枠組みであるCOVAXにも関心が集まりつつある。SDGsの認知度は当初は低かったが、ここ数年高まりつつある。SDGsの広報には広告代理店なども関わっている。そのことの是非は置くとしても、一般の認知度を高めることには大いに貢献している。多くの企業が、その会社広報のパフレット、WEBページ、あるいは商品宣伝にもSDGsを言及するようになってきている。これを後押ししているのはESG（環境、社会、ガバナンス）投資の急拡大である。世界持続的投資連合（GSIA）によれば2020年のESG投資は2018年比15%増の35兆ドルに達した¹。ESG投資たるべき要件はSDGsと概ね一致している。企業にとってSDGsを念頭に置かない事業戦略は困難になりつつある。問われるべきは、各企業の戦略が見せかけではない「本物」であるかどうかである。

SDGsはそれに先立つMDGsが「貧困の半減」や「初等教育の完全普及」など具体的かつ測定可能な8目標で構成されていたのに対し、17目標169項目と複雑化し、Inclusiveness/No one left behind（包摂性／誰一人取り残さない）、Resilience（回復力）などの指標化しにくい概念が使われている。これはMDGsの評価として、貧困削減などの目標は世界全体としては達成されたものの、サブサハラ・アフリカなどの地域が取り残されてしまったという認識から出てきたものである。また、気候変動などの結果として自然災害が

激増しているという状況も踏まえられている。東日本大震災がResilience概念の主流化に影響を与えたこともほぼ間違いないと思われる。

SDGsのスタート時点では、169の項目に対応した指標が特定、確認されておらず、作業が続けられていた。このような状況下、SDGsについて評価の観点から研究を行うべきという長尾眞文国際大学教授（当時）の提唱によって2016年に日本評価学会にSDGs研究部会が立ち上げられた。この準備、基礎固めとして、2015年12月に沖縄のJICA国際センターで開催された第16回全国大会でラウンドテーブル「サステナビリティをどう評価するか」を実施した。ここでは評価で取り扱うべき重点分野の検討、特に分野横断的な分野における評価指標、付加価値ベースの経済指標を使い続けることの問題、民間企業で使われているバランス・スコアカードの活用、などの議論が行われた。

その後2016年度は研究部会活動が本格的に行われ、日本の地域社会の持続性の取り組みなどの報告も含め、SDGsの評価に関する多面的な検討が行われた。また2016年度には、日本学術振興会科学研究費助成事業（JP16K13348）の適用も受けることになった。一連の議論の成果は2017年12月に新潟コンベンションセンターで開催された第18回全国大会で国際セッション「SDGsのためのコミュニティーレベルでの評価能力開発」および共通論題セッション「SDGsの評価」の二つにおいて研究部会での成果報告と、それを発展させた議論を行った。

今回の特集はこの新潟大会での成果と議論をさらに深化させることを目指したものである。大きな論点は三つある。その一つは、キャパシティーに関わるものである。SDGsの評価のためには特に途上国で評価能力をどの様に構築していくかが大きな課題である。Naidoo氏の寄稿、村岡会員・日野氏の論文はこの問題に向き合っている。長尾会員の実践・調査報告は、SDGsの評価をどのようにしてキャパシティー・ディベロップメントに活用することができるか、その取り組みについて論じている。

二番目の課題は、SDGsの理念的・抽象的目標の評価可能性に関するものである。米原会員の論文は、SDG理念型目標の代表例であるSDG4.7のESDに注目し、公立小学校においてESD評価のアクション・リサーチを通じて、ローカルな文脈に適合した「形成的・参加型」評価の必要性を論じている。

三番目の課題は、SDGsが多くの目標を包含することにより、全体の評価セオリーが見えなくなっている問題である。林の論文はSDGs第7目標を通じて、個別目標の指標が必ずしもSDGs全体の持続性を高めることにはつながらないことを指摘している。

研究部会や新潟大会で論じられた上記の論点に加えて、江口会員の研究ノートは改訂されたDAC評価項目とSDGsの関係について示唆に富むものである。極めて時宜を得たものであるので、特集の一つとした。

SDGs中間年が迫る中、SDGsの評価に関して議論を喚起することになれば幸甚である。

注記

【寄稿】

Evaluation Capacities to Advance Sustainable Development for All¹

Indran A. Naidoo

Director, Independent Office of Evaluation
International Fund for Agricultural Development, Rome, Italy²
i.naidoo@ifad.org

Abstract

The Independent Evaluation Office (IEO) of the United Nations Development Programme (UNDP) plays a critical role in advancing the credibility of UNDP's programme support and sustainable development. Providing an overview of IEO's evaluation activities, this paper describes its mandate and approach in supporting accountability within UNDP and in building countries' national evaluation capacities. Specifically, it suggests there is a need for investment in evaluation for greater development effectiveness at the country level. The paper illustrates challenges in linking national evaluation and Agenda 2030, and IEO's initiatives to support countries in diagnosing their evaluation capacities. It suggests the efforts of the United Nations should focus on addressing institutional and process issues to accelerate evaluation capacities. To this end, the United Nations Evaluation Group (UNEG), as an interagency professional network, with its tools and expertise, should undertake advocacy efforts to advance the role of evaluation in the SDG follow up and review process.

Keywords

UNDP, evaluation methods, national evaluation, Agenda 2030, SDGs

1. Introduction

This paper provides an overview of the role of the Independent Evaluation Office (IEO) of the United Nations Development Programme (UNDP) in advancing “critical reflection”, as a “public democratic and accountability good”, to help entrench and advance the credibility of UNDP's programme support. UNDP is the largest development organization within the United Nations (UN) family, has been in field-based operations for over 50 years, and its work spans over 170 countries and through several iterations of Strategic Plans, supports programmes endorsed by the Member States for an ambitious and complex development agenda. The work of the organization is made possible by its historic track record and partnership approach, funded through voluntary contributions, largely by donors and governments, who show confidence by investing in the organization to use its global presence and expertise to advance development. A significant part of the trust and confidence is placed by UNDP in its independent oversight architecture, which includes the independent evaluation unit.

The nature of the work of IEO requires unconstrained access to UNDP programmes for assessment, and the ability to report directly on its findings to the Executive Board,³ programme countries and related programme partners. IEO has a mandate that relates to both accountability and learning, which has generated the highest levels of transparency on organizational performance. In the past decade, over 150 assessments of UNDP were undertaken, providing a high repository of development evaluation assessments to draw upon for accountability and learning purposes, some of which are shared through this paper.

The importance of robust and credible evaluation capacity at the country level is more critical today than during the Millennium Development Goals (MDGs) period given the obligations for country reporting on their Sustainable Development Goals (SDGs) progress. Apart from the work of the largest evaluation office in the UN system, IEO has further contributed to the evaluation profession in its role as Secretariat of the United Nations Evaluation Group (UNEG), and its particular Vice-Chairpersonship of the key stream of evaluation work relating to methodological development and advancement. This includes work on UNEG norms and standards, professionalization, peer reviews, humanitarian, culture and gender evaluation, system-wide evaluation and critically, evaluation of the SDGs. It is in areas such as the SDGs that the IEO has been the most active, using its National Evaluation Capacity (NEC) series to appropriately theme and thus focus global events. IEO has been able to advance the global SDG conversation, given its evaluation leadership role, and also bring in related evaluation issues such as appropriateness of methodologies, evaluation criteria, etc. This paper concludes with one of the IEO's tangible contributions, that is the SDGs Diagnostic Guidance, which is being tested to countries in their reporting on SDG progress as required by the UN.

Following the introductory section, Section 2 of the paper describes the mandate and approach of IEO as it relates to its function of supporting the accountability function within UNDP, and building national evaluation capacities. The section also provides an overview of the key evaluation activities of IEO. Section 3 discusses key issues for evaluations in the context of SDG programming. Section 4 analyses challenges in linking national evaluation and Agenda 2030 and IEO initiatives in supporting countries to diagnose their evaluation capacities. The last section looks at options for accelerating national evaluation capacities and the role of UNEG in supporting this task.

2. The Independent Evaluation Office (IEO) of the UNDP

2.1 Evaluation mandate

The work of IEO, an independent unit reporting directly and accountable to the Executive Board of UNDP, is governed by an Evaluation Policy (UNDP 2019), with key provisions contained therein to ensure structural, budgetary, human resources and evaluation content and dissemination decisions. The accountability function of IEO is strengthened through engagement with the Audit and Evaluation Advisory Committee, whilst methodological quality assurance occurs through the Evaluation Advisory Panel. The evaluation budget for the organization is set at 1% of total resources, with 0.2% dedicated to IEO. The key areas of IEO's work are: (a) conduct Independent Country Program Evaluations (ICPEs) of UNDP country programmes in 170 countries, informing the new country programmes; (b) conduct thematic and corporate evaluations based on UNDP Strategic Plans (UNDP Executive Board 2008 and 2013); (c) quality assess decentralized or self-evaluations and maintain a repository of these; and (d) support efforts to build national evaluation capacity.

As the largest evaluation office in the UN system, IEO provides leadership as Secretariat to the professional body of UN evaluators, the UNEG. The UNEG is a voluntary professional association of 46 United Nations funds,

programmes, specialized agencies and affiliated organizations. IEO has provided leadership to the UNEG in its formation and development, and has consistently provided technical and financial support.⁴ The UNEG strategy for 2014-2019 has four strategic focus areas : (a) evaluation functions and products that uphold UNEG norms and standards; (b) evaluation in support of accountability and programme learning; (c) United Nations system-wide initiatives and emerging demands; and (d) UNEG benefits from contributions to an enhanced global evaluation profession. A key achievement of UNEG has been the inclusion of evaluation in the 2030 Agenda, with the UNEG playing a critical role in advocating the importance of evaluation during intergovernmental negotiations to shape it.

2.2 Scope of UNDP's work

UNDP delivers complex and ambitious development support in areas such as sustainable development, inclusive governance, and enhancing resilience. It is the largest development agency in the United Nations, maintaining a presence in 170 locations across the globe, with dedicated offices in 130 countries working across different development contexts. UNDP, which turned 50 years old in 2017, has a strong and entrenched global presence, and its support is sought to strengthen government capacities across a range of development areas it supports: from addressing issues related to crises, democratic processes, local development, rule of law, to environment management within its broader poverty reduction mandate. It supports efforts in the most difficult and complex development contexts globally, with programmes in countries such as Afghanistan, the Democratic Republic of the Congo and Iraq, addressing a broad mandate, and thriving to meet high expectations from citizens, governments, other development agencies and donors who expect high-quality support. Its budget of USD 15.8 billion for the period 2014-2017, is inadequate for the demands, which have increased as new development issues and crisis have emerged, in particular, environmental, social and political challenges. It has a pro-poor agenda, and to make an impact in this area and be recognized for such is difficult.

2.3 Key activities of IEO

2.3.1 Independent Country Programme Evaluations

A core part of IEO's work relates to assessing the performance of the country programmes towards the end of the cycle, through independent evaluation, as a key input into the next programme. The Independent Country Programme Evaluations (ICPEs), prior to 2017 called Assessment of Development Results (ADRs), capture and present evaluative evidence of UNDP's contributions to development results at the country level, as well as the effectiveness of UNDP's strategy in facilitating and leveraging national efforts to achieve development results. Based on the principle of national ownership, IEO seeks to conduct ICPEs in collaboration with the national government whenever agreed and possible. The ICPEs support greater UNDP accountability to national stakeholders and programmes, serve as a means of quality assurance for UNDP interventions at the country level, are used in the preparation of the new country programmes, and contribute to learning at corporate, regional and country levels.

IEO has completed 106 ICPEs since 2002, which also form the basis for thematic evaluations that are periodically conducted to reflect on the extent to which strategic programme thrusts (poverty reduction, gender, sustainable environment, governance, etc.) are realized. The ICPEs that have been conducted by IEO since 2002 across the five regions of UNDP programme are presented in Table 1.

Table 1 UNDP Programme Countries Covered by the Independent Country Programme Evaluation

Africa		Arab States	Asia and the Pacific		Europe and the CIS	Latin America and the Caribbean	
1. Angola*	17. Madagascar	1. Algeria	1. Afghanistan*	16. Pakistan	1. Albania	1. Argentina	9. Dominican Republic
2. Benin	18. Malawi	2. Djibouti	2. Bangladesh*	17. Papua New Guinea	2. Armenia	2. Barbados	10. Ecuador
3. Botswana	19. Mali	3. Egypt	3. Bhutan*	18. Philippines*	3. Bosnia and Herzegovina*	3. Brazil	11. El Salvador
4. Burkina Faso	20. Mauritania	4. Iraq	4. Cambodia	19. Sri Lanka	4. Bulgaria	4. Chile*	12. Guatemala*
5. Cameroon	21. Mozambique	5. Jordan*	5. China*	20. Thailand	5. Croatia	5. Colombia*	13. Guyana
6. Comoros	22. Namibia	6. Kuwait	6. East Timor	21. Timor-Leste	6. Georgia	6. Costa Rica	14. Honduras
7. Congo Republic*	23. Niger	7. Libya	7. Fiji	22. Viet Nam*	7. Kyrgyzstan	7. Cuba	15. Jamaica*
8. Congo DRC	24. Nigeria	8. Morocco	8. India*		8. Moldova	8.	16. Mexico
9. Côte d'Ivoire	25. Rwanda*	9. Somalia*	9. Indonesia		9. Montenegro		17. Nicaragua
10. Equatorial Guinea	26. Sao Tome and Principe	10. Sudan	10. Lao		10. Serbia		18. Paraguay*
11. Ethiopia*	27. Senegal	11. Syria	11. Malaysia		11. Tajikistan		19. Peru
12. Gabon	28. Seychelles	12. Tunisia*	12. Maldives		12. Turkey*		20. Uruguay
13. Ghana	29. Sierra Leone*	13. United Arab Emirates	13. Mongolia		13. Ukraine		21. Venezuela
14. Kenya	30. Tanzania	14. Yemen*	14. Nepal		14. Uzbekistan		
15. Liberia	31. Togo						
16.	32. Uganda						
	33. Zambia						
	34. Zimbabwe						

128

ADR/ICPS were conducted in total since 2002

66

ADR/ICPE were conducted since 2012

*Countries covered by more than one ICPE/ADR

Note: Since 2012, 102 out of 140 UNDP countries have been assessed by country-level assessment. With IEO's plan for 100% coverage, ICPE will eventually cover all countries

Source: UNDP, IEO *Evaluation Resource Centre* (data as of January 2020) 2.3.2 Global thematic (corporate) evaluations

2.3.2 Global thematic (corporate) evaluations

Global thematic (corporate) evaluations are conducted by IEO to assess UNDP performance in areas that are critical to ensuring sustained contribution to development results in the context of emerging development issues and changing priorities at the global and regional levels. To this end, thematic evaluations cover, for example, UNDP policies, key programme areas, partnerships, programmatic approaches, cooperation modalities, or business models. The thematic evaluations have made an important contribution to informing UNDP's programme strategies. The evaluations conducted by IEO since 2012 are presented in Box 1. IEO also engages in evaluations conducted jointly with other UN organizations to assess system-wide performance.

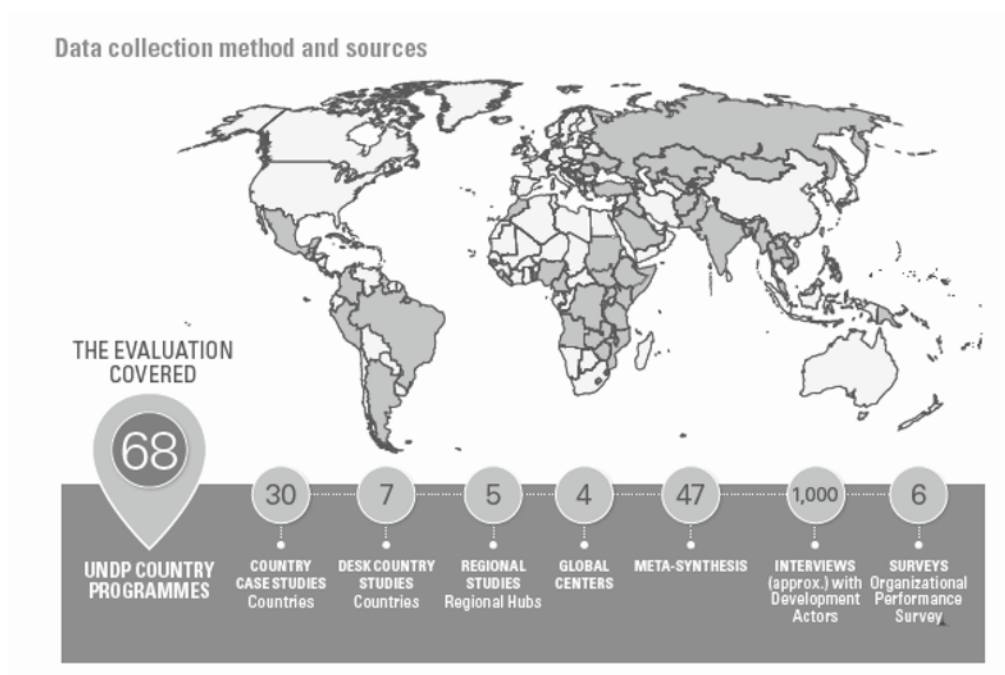
Box 1 Global thematic evaluations conducted by IEO since 2012 ⁵	
Hyperlinks to global thematic evaluations conducted by IEO since 2012	
Evaluation of the UNDP Strategic Plan and Global and Regional Programmes	2017
Joint Assessment of the Institutional Effectiveness of UNDP	2017
Evaluation of UNDP Contribution to Anti-corruption and Addressing Drivers of Corruption	2016
Evaluation of UNDP support to disability-inclusive development	2016
Evaluation of the UNDP contribution to mine action	2015
Joint Evaluation of the Small Grants Programme by GEF/UNDP	2015
Evaluation of the UNDP contribution to gender equality and women's empowerment	2015
Evaluation of the Contribution of the Global and Regional Human Development Reports to Public Policy Processes	2015

Evaluation of the role of UNDP in supporting national achievement of the Millennium Development Goals	2015
Evaluation of UNDP Contribution to South-South and Triangular Cooperation (2008-2011)	2013
Evaluation of UNDP Strategic Plan 2008-2013	2013
Evaluation of UNDP Support to Conflict-Affected Countries in the Context of UN Peace Operations	2013
Evaluation of UNDP Contribution to Poverty Reduction	2013
Evaluation of UNDP contribution to strengthening electoral systems and processes	2012
Evaluation of UNDP Partnership with Global Funds and Philanthropic Foundations	2012

Responding to the signature development solutions of the new Strategic Plan for 2018-2021, IEO will be carrying out six thematic evaluations in the areas of accelerating structural transformations for sustainable development, and enhancing prevention and recovery for resilient societies. In addition, IEO will be assessing UNDP's programme processes, support to multi-donor pool financing and operational services to the UN.

The thematic evaluation uses multiple methods and takes an iterative approach to gathering diverse perspectives in assessing UNDP's performance. Evidence is gathered from document reviews, a meta-analysis of evaluations, regional and country case study missions, interviews and surveys. An example of the multi-method approach used in the recent Strategic Plan evaluation is illustrated in Figure 1.

Figure 1 Multiple-methods used in Strategic Plan evaluation

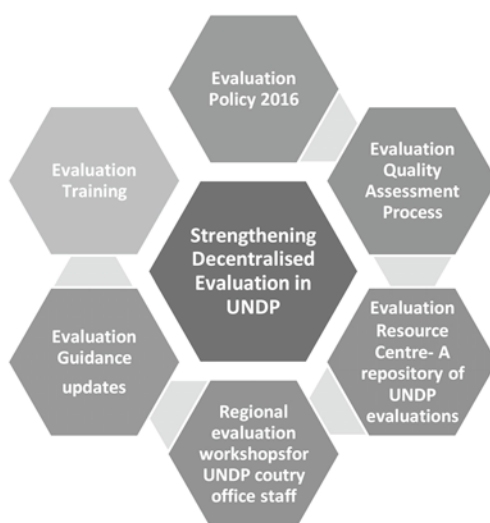


Source: UNDP, IEO (2017), *Evaluation of UNDP Strategic Plan and Global and Regional Programme*

2.3.3 Oversight of decentralized evaluation

The IEO recognizes that high-quality evaluations are critical for results-based management, knowledge generation for wider use, and accountability to programme partners. One of the requirements of the UNDP evaluation policy is that programme units—policy, practice, regional bureaus and country offices—ensure that evaluations inform programme management and contribute to development results (UNDP 2019). There is increased emphasis therefore to strengthen support for decentralized evaluations (those carried out by programme units) to increase the number and coverage of evaluations, improve programme units' compliance with the evaluation policy, enhance the quality of evaluations, and increase policymakers' and stakeholders' use of evaluations. Figure 2 illustrates the activities of IEO to support decentralized evaluation in UNDP.

Figure 2 IEO activities to strengthen decentralized evaluations



Source: UNDP *Independent Evaluation office*

IEO has been undertaking quality assessments of UNDP country office evaluation reports since 2011. A review of the UNDP evaluation policy in 2014 raised concerns about the quality and independence of decentralized evaluations, and made some recommendations to strengthen the ongoing quality assessment process (Le Groupe-conseil baastel ltée 2014). As a result, a hold was put on quality assessment of evaluations in 2015 for two years, while a new policy and a revised strategy to support decentralized evaluations was put in place. The Quality Assessment process is undertaken annually following a rigorous assurance process. The key steps entail:

- **Evaluation Posting** : A programme unit posts an electronic and printable copy of the evaluation report on the Evaluation Resource Centre⁶ as soon as the report is completed. Programme units are responsible for the timely posting and updating of evaluation reports.
- **Verification** : IEO verifies if a report posted on the Evaluation Resource Centre is part of the programme unit's evaluation plan.
- **Quality Assessment** : IEO sends the evaluation report to a contracted reviewer to conduct a quality review.
- **Feedback** : Upon receiving the quality assessment report from the reviewer, IEO reviews the report and then makes it available to the respective programme units.

To ensure quality and consistency, IEO has engaged a pool of evaluation and thematic experts to undertake the quality assessment. These reviewers are senior development professionals with a detailed knowledge of evaluation in their thematic areas, an understanding of UNDP's strategic approach, in general, as well as in evaluation, and also have strong regional and country level knowledge and experience. IEO periodically verifies the quality assessment process to ensure interrater reliability.

Annually, about 250 evaluations conducted by UNDP programme units are quality assessed, including the evaluation of the Global Environment Facility funded projects. For example, of the 255 evaluations assessed in 2016, 28 percent were found satisfactory in that they followed minimum criteria in the Handbook on Planning, Monitoring and Evaluating for Development Results; 47 percent were found to be moderately satisfactory, or came close to meeting expectations, with some minor weaknesses; the rest, 25 percent, were found to be less than satisfactory. This feedback has helped programme units to strengthen the quality of their evaluations.

2.3.4 NEC Conference series: 2009-2017

For over two decades, IEO has supported building national evaluation capacity. One of the key national evaluation capacity development efforts of IEO is the biennial international conferences to provide a knowledge exchange platform on national evaluations. These National Evaluation Capacity Conferences (NEC Conferences) have been taking place since 2009, when the first NEC Conference was held in partnership with the Government of Morocco. Subsequently, the NEC conferences were held in South Africa, Brazil, Thailand, and the recent one in Turkey. The three NEC conferences in Morocco, South Africa, and Brazil discussed the evaluation function, systems, and processes from the perspectives of independence, credibility, and use.⁷ The NEC conferences in Thailand and Turkey addressed evaluation issues pertaining to SDGs. Figure 3 presents the evolution of the NEC Conference series.

Figure 3 Evolution of NEC conferences



Source: UNDP, IEO (2017b). *NEC 2017 Concept Note*.

The fundamental premise of the NEC conferences is to provide the deliberative space for countries with different levels of evaluation capacities to share their lessons and experiences. While providing a regional platform for evaluation practice exchange, the high participation of government representatives makes the NEC conferences

different from other evaluation events. The cutting-edge discussions and workshops on evaluation challenges and how to address them, and lessons from the field, have made NEC conferences an important forum for government actors and evaluation professionals, as well development actors in general.

The NEC conference in Turkey was notable as the largest evaluation knowledge facilitation event bringing together government representatives, evaluation professionals, and civil society actors. Discussions on issues related to “People, Planet, and Progress in the SDG Era” enabled an exchange of lessons, innovations, experiences, and perspectives in developing and adapting national evaluation capacities in light of the SDGs (See Figure 4). The conference discussed ways to respond to SDG specificities and enabled taking stock of some of the evaluation practices that fall short of facilitating SDG evaluations.

Figure 4 Focus of the Istanbul NEC Conference in 2017



Source: UNDP, IEO (2017b). *NEC 2017 Concept Note*.

With over 500 participants from different stakeholder groups from over 110 countries, the NEC Conference provided a platform for partnerships to strengthen national evaluation systems. The high turnout at the event reflects the increasing interest and progress by national governments to develop viable evaluation systems. Discussions such as the relevance of the OECD Development Assistance Committee (DAC) evaluation criteria explored the advantages and limitations of the application of the criteria, and areas where changes are required moving forward with SDG evaluations. The discussion on leaving no one behind addressed areas that need specific attention, for example, gender inequalities, issues of people with disabilities, and regional inequalities in the conduct of evaluation.

3. MDGs to SDGs: What is different this time around for evaluations

While principles such as sustainability, peace, equality, inclusiveness and integrated agenda are not new, the thrust the SDGs give to them provides another opportunity for the countries of the world to change the course of the development discourse in addressing fundamental factors that underpin development and growth. The Agenda stipulates country-led evaluations as an important component of the follow up and review process for the SDGs.⁸

Holistic approach to development programming

The SDGs, unlike their predecessor MDGs, start by acknowledging that the world is more complex and the

focus on social issues, while important, is not by itself sufficient. The SDGs also seek to integrate all the development challenges - inequalities, employment, climate change, governance - to find a more comprehensive response. The lack of clarity on what integrated programming entails remains a challenge both for development agencies as well as governments. A common perception is that integrated programming leads to a loss of focus in individual programme areas. There is limited strategizing for integrated programming, particularly in developing joint outcomes and indicators for designing development interventions. In addition, development funds are allocated to specific programme areas, and government allocation systems are not yet equipped to assign resources and responsibilities for integrated programming. Limited efforts are made to overcome such a compartmentalized approach when addressing complex development issues such as poverty reduction, environment, and disaster risk reduction or climate change adaptation.

National ownership of the evaluation systems and processes

The responsibility of initiating the development processes required to achieve the SDGs by 2030 lies within the states including funding for the SDGs, which is an important shift from the MDGs, and a significant one. While the onus is on the countries, there are bound to be competing priorities, and decisions are bound to be driven by political considerations as well. International development cooperation can play an important role in encouraging and supporting public policy process that underscores key principles of the SDGs.

Universality of the Agenda

The transformative agenda and the universality of the SDGs, their applicability to least developed as well as highly developed countries makes them significantly different from earlier global development efforts. It is a development agenda for all, and NEC-Istanbul reflected this in the most global participation by the government on the evaluation and SDG subject to-date, surpassing all global records. The renewed emphasis on development partnerships for achieving the SDGs is fundamental for collaborative efforts beyond aid. Japan, with an Official Development Assistance (ODA) record of over 60 years, a key funder of multi-lateral assistance, has taken measures to include Agenda 2030 in its development planning. The lessons from Japan's prioritization of eight SDG areas and national reviews will also be critical. Initiatives, such as "SDGs Promotion Headquarters," to foster close cooperation among relevant ministries and government agencies, and to lead the comprehensive and effective implementation of SDG-related measures are critical in facilitating sustainable development oriented partnerships. The success of Japan shall be closely watched in this area, given its strategic prominence as a supporter of a global development agenda in all areas of work across the UN system.

4. Strengthening linkages between national systems, national evaluation and Agenda 2030: the National Evaluation Diagnostic Tool

4.1 Linking National evaluation capacity and Agenda 2030

Countries are interested in evaluation to strengthen their governance agenda and public-sector performance. However, they start from very different levels of both evaluation capacity and general government capacity and resources. Assessments of the national evaluation capacities show that they are evolving, but not at the required pace (Guzman et al. 2014; DIFID 2014; UNDP, IEO 2015a ; Biscaye et al. 2015). There are three broad categories of

evaluation capacity. Mature capacities, where national evaluation is well established, with evaluations mainstreamed across government institutions, with sophisticated data systems. Countries with evolving capacities show a level of commitment to evaluation, but the number of national programmes that are evaluated continues to be limited. Data systems generate information on a set of indicators. Emergent evaluation capacities are in the process of creating an enabling environment for evaluation and only generate basic data. Most countries represent evolving and emergent capacities, with examples of sophisticated evaluation systems remaining an exception.

The deliberations of the NEC Conferences over the years point to constraints in the use of evaluations. There is poor alignment of evaluation systems and national planning processes. Timing and scope of evaluations are often not aligned with planning and budget cycles. Learning loops or formal processes for using/applying evaluation findings is an area that needs attention. Evaluation policies and legislation, while found in several countries, are not always followed up by independent institutional measures.

While national evaluation policies are critical for strengthening evaluation systems and capacities, the experience of the MDG period has pointed to the importance of robust statistical systems for managing development progress (UNDP, IEO 2015a; UNDP, IEO 2015b; UN Joint Inspection Unit 2016). Significant development data challenges remain, and official data is often used for reporting and not so much for public policy decisions or performance monitoring. Enabling a policy environment for national evaluation capacity development is critical for moving forward on SDG-evaluations. Countries are yet to prioritize data capacities (UNDP, IEO 2015b; Dunning & Karlow 2016).

Developing evaluation capacities is evolutionary and seldom linear, and can easily be derailed by political changes. It is often seen as an additional burden and not adequately integrated, and, if so, it is often not given the space for critical judgment due to inadequate evaluation policy. Countries with nascent evaluation functions cannot be expected to produce and use high-quality, cross-sectoral real-time evaluations within a three-, five- or even ten-year time span. The setting up of evaluation systems needs to take into account several complex organizational and capacity issues, most of which are not in place, and which, thus, result in a mismatch between expectations and the actual capacity of the prevailing administrative systems and culture.

4.2 Responding to evaluation capacity needs: the National Evaluation Diagnostic Tool

At the country level, there is considerable demand for tools and resources to strengthen national evaluation frameworks for the SDGs and inform country-led evaluation processes and systems that are central to the follow-up and review of the progress on SDGs. Assessing national evaluation needs and capacities are critical as countries are integrating the SDGs into their national development strategies. Responding to this demand, IEO developed the National Evaluation Guidance and an Online self-assessment tool (UNDP, IEO 2019).

The Online Self-Assessment Tool for national evaluation diagnostics and strategizing, developed by IEO, provides tools for step-by-step evaluation capacity diagnosis, needs assessment, and action points in developing a country's evaluation framework for the national development strategies (which includes the SDGs). The Self-Assessment Tool considers the development of national evaluation capacities in the era of the SDGs an effort to integrate, strengthen, and connect three interlocking elements, i.e., national systems of the government, Agenda 2030, and evaluation systems and processes. What is also fundamental to the IEO approach is that national evaluation capacity cannot be seen in isolation from other government systems and processes.

The Self-Assessment Tool provides a flexible and practical framework to enable governments and other development actors to develop a systematic approach to determining key areas, pathways, and parameters for evaluating national development strategies and SDGs. It also details action points to be considered in developing a country's evaluation framework for the national development strategies and the SDGs, and will facilitate identifying

various steps in developing a country-level evaluation framework. The online tool can be tailored for use by federal government, regional/state government as well as local levels. The four modules of the Self-Assessment Tool are presented in Table 2. The modules enable an understanding and building of an enabling environment for national evaluation systems, linking evaluation systems and national and local SDG (development) planning; diagnose existing evaluation capacities; strengthen and institutionalise evaluation capacities; and integrate SDGs specificities into evaluation processes and practices.

Table 2 Module of the Self-Assessment Tool

<i>Module 1</i>	<i>Module 2</i>	<i>Module 3</i>	<i>Module 4</i>
Understand enabling environment for national evaluation systems	Connecting national systems and Agenda 2030 to the national evaluation system	Strengthening and institutionalising evaluation capacities	Integrating SDGs / sustainable development specificities in evaluation processes and evaluations

Source: UNDP, IEO (2019)

The Guidance and the Self-Assessment Tool situates evaluative practices within the context of other public-sector feedback mechanisms and recognizes that evaluation is a practice that is built over time, rather than a set of outputs. It is based on the principles that evaluations will be country-led, take an integrated view of systems and development, and will be learning and utilization-focused. It also incorporates the principles of the inclusive and participatory approach to evaluations, emphasizing the importance of integration into national systems and not creating additional burdens or duplicating systems.

A set of activities will be carried out by IEO for the dissemination of the Tool at the country level and to evaluation and development stakeholders. The activities include those aimed at increasing the awareness about the Tool, enabling its rollout, orientation and training to relevant stakeholders who are likely to use the Tool, and initiatives for wider reach to evaluation and development stakeholders. The Tool will be rolled out initially in 10 countries across five regions, and at the regional workshops where there will be an opportunity to engage with a higher number of countries willing to champion the Tool. IEO recognizes that the Tool complements other ongoing efforts such as Mainstream, Acceleration and Policy Support (MAPS), sector capacity assessments, High-level Political Forum on Sustainable Development (HLPF) related events, where the synergies of the Tool can provide a holistic support on SDG-mainstreaming and analytics.

5. Need for joint action: the role of UNEG

There is need for considerable investment in evaluation for greater development effectiveness at the country level. While this needs to be country-led, efforts at multiple levels are critical to support national efforts to strengthen evaluation processes and systems. A more systemic approach to national evaluations is needed, which entails a holistic view of development and looking at evaluations as an interaction between a complex set of actors and actions. Efforts of the UN, as well as other actors, should, therefore, focus on addressing institutional and process issues that can accelerate evaluation capacities.

At the UN, the SDGs require the attention of several agencies and actors. Each entity has its mandate as well as comparative advantages, rich experiences and potential opportunities. UNEG, as an interagency professional network, is supporting this work by providing tools and technical skills. Evaluation of SDGs is the dedicated work of

UNEG, which is well-placed to undertake advocacy regarding the role of evaluation in the SDG follow-up and review process. The UNEG “Norms and Standards” and its “Competency Framework” recognize the importance of fostering national ownership and capacity for evaluation (UNEG 2016a). UNEG is also engaged in efforts on promoting a credible independent system-wide evaluation mechanism. UNEG is exploring ways to determine the role of UN organizations in the 2030 Agenda follow-up and review process at different levels, and the role of UNEG in contributing to that. Evaluation reforms are needed for UN agencies to respond to SDG-related programme support appropriately (UNEG 2016b). Partnership with forums such as the Evaluation Capacity Group (ECG) and the OECD DAC evaluation network has been important in this process.

The national evaluation capacity initiatives need to be more regular, focused and deeper drawing in the oversight units of the government into the conversation. Countries use a variety of accountability and transparency mechanisms, and there is need for an alignment between such complementary systems. Whereas the demand for evaluation as a feedback loop is rising, monitoring nonetheless remains a much more prominent practice, especially at the level of many policy-makers. There remains a certain level of apprehensiveness to the notion of ‘evaluation’ – with its judgemental and diagnostic element about past performance. This often reflects the lack of a culture where critical reflection is encouraged, when it should be more important to “improve” government performance, than “prove” performance, noting the inherent positive bias that comes from all forms of self-reporting by evaluands. Complementary accountability mechanisms, those that are already well established, can be entry points to strengthen the evaluation culture.

SDG success is a combination of political commitment and technical competence, genuine and sustainable engagement with civil society, academia and communities so that it is localized to the extent it makes a tangible impact. What countries are likely to achieve over the course of the next five, ten, and fifteen years hinges on their current points of departure. Agenda 2030 puts heavy emphasis on national ownership and leadership, in line with the Paris Declaration on Aid Effectiveness and the Accra Agenda for Action.⁹

The inclusive and participatory process of the SDGs resulted in a comprehensive, far-reaching and people-centred set of universal and transformative goals and targets, perhaps the global community’s single largest developmental agenda, roadmap and hope for a more equitable, inclusive, resilient, peaceful and sustainable planet. Ownership is broad-based and strengthens the demand for accountability for results. This is the context to which national evaluation practice and capacities must respond.¹⁰ Agenda 2030’s follow-up and Review Principles provide strong backing for evidence-based decision-making; shift the attention from the MDG focus on reporting at the global level to an SDG focus of learning and improvement at the local level; and mandate the UN system to support countries in the development of monitoring and evaluation systems. All evaluations conducted must include SDG lens and need to be a reflection on what constitutes appropriate criteria for development outcomes.

Notes

- 1 An earlier version of this paper was presented at the Annual Conference of the Japan Evaluation Society held in Niigata, Japan in December 2017.
- 2 Former affiliation: UNDP, Indran Naidoo, Director, Independent Evaluation Office, United Nations Development Programme, New York, USA
- 3 The Executive Board of UNDP, UNFPA and UNOPS, comprised of 36 members, was created by General Assembly resolution 48/162 of 20 December 1993. The Executive Board superseded the 48-member Governing Council on 1 January 1994.
- 4 See Annual Reports on the Evaluation for the Years 2015, 2016, 2017, <http://web.undp.org/evaluation/annual-report/are.shtml>
- 5 The global thematic evaluations conducted by IEO can be found here: <http://web.undp.org/evaluation/evaluations/>

thematic-evaluations.shtml

6 See Evaluation Resource Centre, <http://erc.undp.org>

7 See NEC Morocco <http://www.nec2013.org/downloads/NEC-2009-Proceedings.pdf>; NEC South Africa <http://www.nec2013.org/downloads/NEC-2011-Proceedings.pdf>; <http://www.nec2013.org/>; NEC Brazil <http://www.nec2013.org/>; NEC Thailand <http://www.nec2015.net/>; NEC Turkey <http://web.undp.org/evaluation/nec2017/overview.shtm>

8 UN General Assembly, 2016. "A/RES/70/299 - Follow-up and review of the 2030 Agenda for Sustainable Development at the global level" (2016); UN ECOSOC. "E/HLS/2016/1 - Ministerial declaration of the 2016 HLPF" .

9 Paris Declaration on Aid Effectiveness and the Accra Agenda for Action, 2005. <http://www.oecd.org/dac/effectiveness/parisdeclarationandaccraagendaforaction.htm> ;

10 Key note address by Michelle Gyles-McDonnough, Director, Sustainable Development Unit, Executive Office of the Secretary-General, at the 5Th Global National Evaluation Capacities Conference, on 'People, Planet and Progress in the SDG Era' , Istanbul, 18 October 2017, See <http://web.undp.org/evaluation/nec2017/>

References

Biscaye, P., LaFayette, M., Martin, A., Richardson, M, True, Z., and Anderson, C (2015). *Evaluating Country-Level Government Monitoring & Evaluation Systems*. University of Washington, Evans School of Public Policy & Research Group, Seattle.

DIFID (2014). *Demand for and Supply of Evaluations in Selected Sub-Saharan African Countries*, London. <https://www.gov.uk/dfid-research-outputs/demand-for-and-supply-of-evaluations-in-selected-sub-saharan-african-countries>

Dunning, C. & Kalow, J. (2016). *SDG Indicators: Serious Gaps Abound in Data Availability*, Washington, DC: Center for Global Development. <http://www.cgdev.org/blog/sdg-indicators-serious-gaps-abound-data-availability>

Guzman, M., Irarrazaval, I., and de los Rios, B. (2014). *Monitoring and Evaluation System : The Case of Chile 1990-2014*. Evaluation Capacity Development (ECD) working paper series; no. 29. World Bank., Washington, DC. <https://openknowledge.worldbank.org/handle/10986/21313> License: CC BY 3.0 IGO.

Joint Inspectin Unit (2016).

Le Groupe-conseil baastel ltée (2014). *Review of the UNDP Evaluation Policy*, Gatineau, Canada. http://web.undp.org/evaluation/documents/policy/review/Baasterl-UNDP-evaluation_PR_2014.pdf

UNDP (2019). *The Evaluation Policy of UNDP*, (DP/2019/29), New York.

http://web.undp.org/evaluation/documents/policy/2019/DP_2019_29_E.pdf

UNDP Executive Board (2008). *UNDP strategic plan, 2008–2011 – Accelerating global progress on human development*, (DP/2007/43/Rev.1), New York, May. See, <http://web.undp.org/execbrd/pdf/dp07-43Rev1.pdf>

UNDP Executive Board (2013). *UNDP strategic plan, 2014–2017 – Changing with the World*, (DP/2013/12), New York, June. See http://www.undp.org/content/dam/undp/library/corporate/UNDP_strategic-plan_14-17_v9_web.pdf

UNDP, IEO (2015a). *Evaluation of UNDPs support to MDGs*, New York.

UNDP, IEO (2015b). *Evaluation of Global and Regional Human Development reports*, New York.

UNDP, IEO (2017a). *Evaluation of UNDPs Strategic Plan, Global and Regional Programme*, New York. <http://web.undp.org/evaluation/evaluations/thematic/sgrp.shtml>

UNDP, IEO (2017b). *NEC 2017 Concept Note*. See <http://web.undp.org/evaluation/nec2017/>

UNDP, IEO (2019). *National Evaluation in the SDG Era: Online Self-Assessment Tool for National Evaluation Diagnostics and Strategizing*, New York. <http://web.undp.org/evaluation/evaluations/documents/diagnostic-tool/National%20Evaluation%20Diagnostic%20Guidance.pdf>

UNEG (2016a). *UNEG Norms and Standards for Evaluation*, New York <http://www.uneval.org/document/detail/1914>

UNEG (2016b). *Evaluation in the SDG era: lessons, challenges and opportunities for UNEG*, Volume 2, New York.

UN Joint Inspection Unit (2016). *Evaluation of the contribution of the United Nations development system to strengthening national capacities for statistical analysis and data collection to support the achievement of the Millennium Development Goals (MDGs) and other internationally-agreed development goals*. (JIU/REP/2016/5), New York.

List of links to UNDP reports

UNDP, IEO Evaluation Resource Centre

Evaluation of the UNDP Strategic Plan and Global and Regional Programmes

Joint Assessment of the Institutional Effectiveness of UNDP

Evaluation of UNDP Contribution to Anti-corruption and Addressing Drivers of Corruption

Evaluation of UNDP support to disability-inclusive development

Evaluation of the UNDP contribution to mine action

Joint Evaluation of the Small Grants Programme by GEF/UNDP

Evaluation of the UNDP contribution to gender equality and women's empowerment

Evaluation of the Contribution of the Global and Regional Human Development Reports to Public Policy Processes

Evaluation of the role of UNDP in supporting national achievement of the Millennium Development Goals

Evaluation of UNDP Contribution to South-South and Triangular Cooperation (2008-2011)

Evaluation of UNDP Strategic Plan 2008-2013

Evaluation of UNDP Support to Conflict-Affected Countries in the Context of UN Peace Operations

Evaluation of UNDP Contribution to Poverty Reduction

Evaluation of UNDP contribution to strengthening electoral systems and processes

Evaluation of UNDP Partnership with Global Funds and Philanthropic Foundations

UNDP, IEO (2017), Evaluation of UNDP Strategic Plan and Global and Regional Programme

UNDP Independent Evaluation office

UNDP, IEO (2017). NEC 2017 Concept Note.

(Accepted 31st July, 2021)

【研究論文】

SDGs の時代における評価能力構築

村岡 敬一

外務省大臣官房 ODA評価室
DAC開発評価ネットワーク副議長
Muraoka.Keiichi2@jica.go.jp

日野 類子

外務省大臣官房 ODA 評価室
hino.ruiko@gmail.com

（報告者の所属と肩書は本稿執筆時（2018年4月）のものであり、現在は村岡が独立行政法人国際協力機構、日野が一般財団法人国際開発機構に所属している）。

要 約

本稿は、国際的潮流、我が国のODAを通じた取り組み、開発途上国側の現状と課題の検討を通じて、SDGsの時代における評価能力構築を論じたものである。2030アジェンダの「誰一人取り残さない」理念とそれに基づくSDGsは途上国側の評価能力構築に加え、参加型評価等新しい評価手法導入の必要性を喚起するものであり、「課題」であると同時に「機会」でもある。開発途上国側で必要とされる評価能力構築に関する支援のニーズは膨大であり、開発途上国にとってSDGsの時代が「機会」となり得るためには、我が国を含むドナーの支援が不可欠であるといえる。

キーワード

SDGs、評価能力構築、外務省、ODA、DAC開発評価ネットワーク

1. はじめに —SDGsと評価: 2030開発アジェンダから—

2015年9月の国連特別総会で採択された「持続的な開発のための2030アジェンダ（以下、2030アジェンダ）はその中で持続可能な開発目標（SDGs）を掲げるとともに、終章に「フォローアップとレビュー」に関する提言が設けられている。2030アジェンダでは「誰一人取り残さない（leave no one behind）」を理念として掲げ、具体的な指標及びフォローアップとレビューメカニズムの策定は国連経済社会理事会のもとで開催される国連ハイレベル政治フォーラム（High-level Political Forum, HLPF）¹においてモニタリングさ

れることになる。

一方、SDGsのフォローアップとレビューは開発協力に携わる評価関係者にとっても大きな関心事となっており、世界銀行や国連開発計画（UNDP）においても評価のあり方や開発途上国における評価能力の構築が議論されている。こうした議論を受けて経済協力開発機構・開発援助委員会（The Organisation for Economic Co-operation and Development's Development Assistance Committee, OECD/DAC）では2017年10月に採択されたハイレベル・コミュニケ²に基づき、SDGs時代に即した評価項目（クライテリア）の見直し作業を開始し、2019年12月に新たな評価基準が採択されている³。SDGs評価における課題については、評価クライ

テリアを含む評価手法の課題、評価能力の課題、統計・データの信憑性及び統計制度の課題等が挙げられるが、本稿ではSDGsの時代における評価能力構築について、国際的潮流、我が国のODAを通じた取り組み、開発途上国側の現状と課題の検討を通じて、SDGsの評価に向けた評価能力構築課題、その背景・理由、必要となる対応について具体的に論じる。

2. 2030アジェンダにおける評価の位置づけ

2030アジェンダ（国際連合2015）の最終章「フォローアップとレビュー」には次の記載がある。

74.（基本原則）

すべてのレベルにおけるフォローアップとレビュープロセスは、次の原則によって導かれる。

*g.*これらは、各国の主導で行われる評価やデータに基づく正確で根拠のあるものである。各国が行う評価やデータは、高品質で、アクセス可能、時宜を得た、細分化されたデータに基づくものであり、具体的には、収入、性別、年齢、人種、民族的属性、移住者の法律上の地位、障害、地理的属性及びその他各々の国内での状況に関連ある特徴等を踏まえたデータである。

*h.*これらは、特に、アフリカ諸国、後発開発途上国、小島嶼開発途上国、内陸開発途上国、中所得国等の開発途上国における国家資料システム及び評価事業の強化を含む能力開発の拡大を必要とする。

ここで、開発途上国の評価能力構築に関する記述として特に重要な記述は、*h*項の「開発途上国における国家資料システム及び評価事業の強化を含む能力開発の拡大を必要とする」であるといえる。

3. 開発途上国の評価能力構築に向けた評価関係者の動き

既述のとおり2030アジェンダのフォローアップとレビューはHLPFにおいて進捗状況を確認することになっており、2017年のHLPFでは我が国を含む43か国が自発的国別レビュー（Voluntary National Review, VNR）を行った⁴。右に報告された開発途上国の評価能力構築の現状については本稿後段にて考察するが、ここでは国連経済社会理事会での議論を補完しつつ繰り広げられた開発途上国の評価能力構築に関する国際的な動きを概観する。

（1）援助効果向上に関するパリ宣言からグローバルパートナーシップへ

開発途上国の評価能力構築のテーマは2005年2月にOECD/DACによりパリで開催された「第2回援助効果向上に関するハイレベルフォーラム」において採択された「援助効果向上に関するパリ宣言」において広く認知された。同宣言では援助効果向上の5原則のひとつに「開発成果管理」を掲げ、開発途上国の計画、予算管理、評価等の援助実施管理に関する制度を強化し、相互の連関性を高めることにより、開発の成果を高めることを謳っている（外務省2007、p.33）。これと前後してDACの下部組織である開発評価ネットワーク（EvalNet）ではEvaluation Capacity Development（ECD）を審議テーマに加え、ODAを通じた途上国の評価能力構築支援について議論している。その後、パリ宣言の精神は「効果的な開発協力に関するグローバルパートナーシップ（GPEDC）」に受け継がれ、爾来、我が国においても途上国の評価能力の構築に向けた協力が外務省、JICAを通じて行われてきた。なお我が国外務省がODA評価ワークショップを開始したのは2001年であり、評価の分野においても開発途上国の制度構築に早くから着目していたことが窺われる。外務省は2005年に開催されたEvalNet会合においてこうした取り組みを報告している。

(2) アジェンダ2030の採択、SDGsの合意

そうした中、2015年の国連特別総会におけるアジェンダ2030の採択とSDGsの合意は、途上国の評価能力構築の重要性に新たな光を当てる契機となった。国連ミレニアム開発目標（MDGs）の終盤を迎え、SDGsの合意に向けた交渉が国連においてすすめられていた2013年9月に国連開発計画（UNDP）がブラジル政府他との共催によりサンパウロで開催した第3回評価能力構築に関する国際会議（NEC）⁵では2015年を「国際評価年」とすることが提唱された。これを受け2014年12月の第69回国連総会において途上国の評価能力構築の重要性とそれに対する国際的な支援を呼びかけた国連決議（United Nations 2015）が採択され、この中で2015年国際評価年が認知された。こうした途上国の評価能力構築にむけた国際的な潮流の背景にはSDGsを見据えた開発政策の策定にあたっては開発途上国自身によるモニタリング・評価システムを経て確認された根拠に基づいているべきであり、そのためには途上国自身の評価能力を向上していくことが重要であるとの認識があったとされる（外務省2016a、p.2）。

(3) 近年の評価に関する国連機関の動向

2015年国際評価年の関連の記念行事が各地で展開される中、UNDPは同年10月に国際評価協会（The International Development Evaluation Association, IDEAS）他との共催で第4回評価能力構築に関する国際会議（NEC）⁶をタイで開催し、SDGs時代における評価原則を謳ったバンコク宣言を採択した。バンコク宣言では前述の国連総会決議に言及しつつ、政府、二国間及び多国間援助機関は途上国の評価能力への支援を開発協力の優先課題として位置づけるよう呼びかけている。

翌2016年3月に国連評価グループ（UNEG）とフォード財団がニューヨークで開催したSDGs評価に関する国際会議には、日本から外務省⁷、JICAの他、本邦NPOが参加している。同会議におけるテクニカルセミナーでは開発途上国の評価能力構築の課題として以下の点がハイライトされた。

- － 2030アジェンダにおいて政治的に合意された「誰一人取り残さない」理念に基づいた開発の成果を評価するためには、途上国側の統計基盤の向上と共に評価能力の向上が不可欠。このためには南南協力、地域内協力も有益である。
- － SDGsにおける評価においては、既存の評価方法にとらわれないタイムリーかつ柔軟性のあるモニタリング・評価を行う必要がある。また評価において「誰一人取り残さない」ためには社会的弱者や少数者を含む開発事業の最終受益者へのプラスとマイナスの影響（インパクト）を測定することが求められるが、そのためにはステークホルダーの声をひろく汲み上げることが出来る参加型評価の手法が有効である。

(4) 近年の開発評価ネットワーク（EvalNet）の動向

一方、OECD/DACの下部組織である開発評価ネットワーク（EvalNet）では2016年4月に開催された第19回会合⁸において開発途上国の評価能力開発に向けた支援強化の方策が議論され、ODA評価ワークショップを通じアジア・大洋州地域の評価能力の向上に一定の成果を挙げてきた日本の経験・知見が改めて注目された。その後、EvalNetでは関心国からなる評価能力開発（ECD）タスクフォースが設置され、2030年に向けた評価能力構築のヴィジョンと使命が翌年2月の第20回会合⁹に報告されている。

4. 外務省主催ODA評価ワークショップの開催

我が国外務省は、記述の通りアジア大洋州諸国の評価能力構築を目的として政府関係者や専門家を招いたODA評価ワークショップを2001年から開催している。その目的はODA評価の手法や課題について、アジア・太平洋地域における理解を増進し、特に途上国の評価能力を向上させること、途上国関係者自身のODA評価能力の向上を通じてドナー国側の援助効率化だけではな

く、途上国側のオーナーシップ・透明性の向上は開発の効率化を目指すものとされている（外務省2017、p.3）。ワークショップでは途上国における組織的なモニタリング・評価体制構築の課題や評価人材の育成の課題が議論されると共に、我が国ODA評価の事例や統計学的手法を用いたインパクト評価や人類学的手法を用いたプロジェクト・エスノロジー評価の事例等も紹介されている¹⁰。またワークショップでは2012年のアジア太平洋評価協会（APEA）の設立を後押しし、その後は域内の評価専門家の招聘等を通じAPEAの活動を側面的に支援してきた。ここではSDGsの合意以降に開催されたODA評価ワークショップでの発表及びその議論を通じてアジア大洋州諸国における評価能力構築の取り組み及び課題について考察する（なお、SDGs合意直後に開催された第12回ODA評価ワークショップではSDGsに関する議論を行っていないため、検討の対象にはしない）。

（1）第13回ODA評価ワークショップ（2015年 東京会合）

外務省ODA評価室は2015年国際評価年記念行事の一環として、アジア・大洋州諸国の評価能力構築を目的とした第13回ODA評価ワークショップをJICA及びAPEAの協力を得て2015年11月に東京で開催した¹¹。同会合は国際評価年の記念行事として登録され、域内27カ国、3国際機関が参加した。2日間のワークショップでは「SDGsの達成に向けた評価の新たな課題」と題したセッションを設け、インドラン・ナイドー UNDP独立評価部長及び源由理子明治大学教授・日本評価学会副会長がそれぞれ報告した（報告者の肩書はいずれも当時のもの）。

ナイドー報告は、「SDGsの達成に向けたUNDPの役割」と題する報告の中で、戦略的なアプローチとあらゆるステークホルダーの参加、そして評価を最大限に活用することが、SDGsの目標達成のために重要であることを強調した。続いて、源報告は、「SDGsにおける変革型評価の適用とその課題」と題する報告においてSDGsの包括性あるいは統合的アプローチには、持続的開発、社会的に脆弱な立場にある人々との連帯、民主

的統治そして全ての関係者の参加等の概念が含まれており、評価情報の潜在的な活用者は政府関係者のみならず民間企業や市民社会も含まれるとした。そしてその開発の過程では多様なアクター間の相互作用的な状況を常に生み出していくインターアクティブ・ガバナンスへの変容が求められるのではないかとし、変革型の社会的介入を評価するアプローチとしてプログラム評価理論と参加型評価のアイデアを取り入れた変革型評価を用いることを提案した。

両報告を受けて、アジア大洋州諸国参加者間で、SDGsの体制下での評価における国際協力、特に南南協力¹²を強化することの必要性等について活発な議論が展開された。

（2）第14回ODA評価ワークショップ（2016年 ハノイ会合）

翌2016年11月にベトナム・ハノイで開催された第14回ODA評価ワークショップはアフリカを含む32カ国が参加した初のAPEA国際評価会合の一環として実施された¹³。ここではフィリピン国家経済開発庁（NEDA）と在フィリピン日本大使館が合同で実施した評価の結果が報告された（外務省2016b）。この合同評価は、外務省ODA評価室による「被援助国政府・機関等による評価」スキームを活用して、防災分野の日本のODAを評価対象として実施されたものである。同報告に続く議論では、合同評価が途上国の評価能力の向上と共にドナー・パートナー双方の相互説明責任の向上に資することが確認された。

（3）第15回ODA評価ワークショップ（2018年 コロンボ会合）

2018年2月にスリランカ・コロンボで実施された第15回ODA評価ワークショップでは、「持続可能な開発目標（SDGs）の時代における評価能力開発」という副題の下に4つのセッションの中で6つの報告が行われ、参加者間で議論がなされた¹⁴。ここでは、第1セッション：「ODAにおける評価の役割と国際的な潮流」から、ハウゲ国際連合開発計画（UNDP）独立評価部次長報告（以降ハウゲ報告）とヴィジャヤラガバンアジア開発銀行（ADB）独立評価部課題・国別課上級評価専

門官報告（以降ヴィジャヤラガバン報告）、第2セッション：「SDGsへの対応と取組－インドネシアにおける取組とその課題－」から、石田洋子広島大学教育開発研究センター教授報告（以降石田報告）について概観する。

ハウゲ報告は、「持続可能な開発のための2030年アジェンダ」の構造について詳細な説明を加えた上で、同アジェンダの下での評価の「課題」と「機会」について論じた。

「課題」として挙げられた点は、①目標の不明瞭さ、②目標の複雑さ（相互関連性）、③測定可能性（指標の不備）、④成果と要因の因果関係の4点である。SDGsでは、17の目標の下に169のターゲットが設定されている。17の目標は相互に関連しているため、ターゲットの下に設定される指標もまた相互に関連し複雑なものとなる。目標の達成度合いを測るためには、指標が設定され測定される必要があるが、300以上ある指標の中には、各国が保有していないデータや、指標が明確に定義されていないものもある。また指標を測定するには、多くの利害関係者（国外、国内、国内の地方部）との調整が必要となり、刻々と変化する状況への対応も求められる。また、目標の達成（成果）を導く要因も変化するため、同要因を特定することも難しさを伴う。

「機会」としては、①SDGsレビュー枠組の強化、②国家評価能力の開発、③革新的な技術の利用、④情報へのアクセスの改善が指摘された。ヴィジャヤラガバン報告でも言及されたが、MDGsと異なり、SDGsでは評価の重要性が明示されており、その点において、SDGsとその評価の時代は各国でのレビュー枠組みの強化、ひいては国家評価能力の開発につながる「機会」と捉えられよう。また目標の達成度合いを測るために、革新的な技術、例えば途上国の公的な統計システムの不備を補完すべくビッグデータやリモートセンシング技術による地理情報の活用等の導入が積極的にすすめられる可能性もある。

ヴィジャヤラガバン報告においても、ハウゲ報告と同様SDGsが評価のための新たな機会と捉えられることが指摘された。それに加えて、SDGsは国際的評価コミュニティに対し、目標の達成において、何が機能するのか／何が機能し

ないのか／国・地域・実施機関間で（結果が異なるのはなぜか、その疑問についてのより良い学習が求められること、そして同コミュニティが評価に関する付加価値を高めるきっかけとなる点が示された。

ハウゲ報告でも、国連に提出されるSDGsのVNRにおいて評価機能が十分認知されていない点について問題提起がなされ、VNRガイダンスにおける評価ツールや技法が今後改善されていく予定であることが示された。このように、国際機関を含む評価コミュニティにおいても、SDGsとその評価の時代において、それぞれの組織が付加価値を高めていくことが求められている。同時に各国では評価の「ローカライズ」、すなわち各国の状況に即した評価戦略の策定、実施が求められる。この点は、ハウゲ報告、ヴィジャヤラガバン報告に続く参加者間の議論で確認された点である。

石田報告は、インドネシアにおける評価の「ローカライズ」の状況について報告をした。より具体的には、JICAがインドネシアで実施中の「持続可能な開発目標（SDGs）の計画・運営促進に係る情報収集・確認調査」からの発見・気づきの点として、同国政府の国家の制度構造、グローバル指標の国・地域レベル指標での代替化、仕組みと手順、モニタリングと評価における課題が示された。

インドネシアは、自国の社会経済発展を監視するために必要な69のインドネシア固有の指標をグローバル指標に加えており、SDGsで掲げられる17の目標の下に319の指標を設定している。インドネシアにおけるモニタリング・評価の能力開発に考慮すべき大きな課題として、①SDGs指標と国家中期開発計画プログラムの成果との因果関係の明確化と判断の方法、②地方自治体におけるモニタリングとフィードバック方法、③非国家団体、慈善団体のモニタリング及び協働の方法の3点が指摘された。

同報告に続く参加者間の議論では、インドネシア国家開発企画庁参加者より同国の地方自治体は数が多く、評価能力には一様でないとの指摘があり、地方部におけるモニタリング評価の能力開発の必要性が確認された。

5. アジア諸国のSDGsのフォローアップ・レビューメカニズムの現状と課題

(1) 参加者アンケート・報告から

前述した2018年のコロンボ会合では、招聘した13か国の参加者（表1参照）に対し、事前にSDGsのフォローアップ・レビューメカニズムの現状と課題に関するアンケートを送付し、8か国¹⁵から回答を得た¹⁶（表2参照）。

表1 コロンボ会合参加国

コロンボ会合参加 13 か国	インド、インドネシア、スリランカ、タイ（*）、ネパール（*）、パキスタン（*）、バングラデシュ（*）、フィリピン（*）、ブータン（*）、マレーシア、ミャンマー、モルディブ（*）、モンゴル（*）
* アンケート回答国（8 か国）、下線：VNR 作成国（8 か国、2018 年 3 月時点）	

（出所）筆者（日野）作成。

アンケートに示された回答から、アジア諸国のSDGsフォローアップ・レビューメカニズムの現状と課題に関する傾向について検討する。

まず、フォローアップ・レビューのデザイン、枠組みづくりは、多少の進度の違いはあるがいずれの国においても、すすめられている。自国特有の状況に応じた指標枠組みについても各国の取り組みがすすんでおり、SDGsのローカライズの必要性は認識され、被援助国側にもその需要はあることが分かる。次に、指標測定に必要とされるデータ収集方法の確立についても一定程度の進展がみられる。だが、それに続く実質的なデータの集積・集計システムの確立については多くの国が構築中と回答している。関連省庁、国民へのレビュー結果のフィードバック体制についてはいずれも構築途上にある国が大多数であり、まずはモニタリング・レビュー体制の仕組みづくりが優先的にすすめられているといえる。

(2) 自発的国別レビュー（VNR）から

次に、HLPFのウェブサイトより取得した各国の自主的国別レビュー（VNR¹⁷）についての検討

を行う。この検討を行うのは、前述のアンケート回答者の多くは、SDGsのフォローアップ・レビューを行う役割を負う行政機関（首相府、外務省、国家計画委員会等）では無い組織や部署に属する行政官¹⁸であり、彼らの回答に示された情報の確からしさについては一定程度の留保を付ける必要があり、異なる情報源からの検討が必要と考えられるためである。

参加13か国のうち、VNRを作成しているのは、8か国である。これら8か国のVNRから得られた、指標の設定、データの集積・集計、国家開発政策へのSDGsの統合、フォローアップ・レビューの課題に関する情報を取りまとめた（表3参照）。

表3に示すように、各国のVNRには指標の設定については2か国を除き記載があり、5か国はフォローアップ・レビューメカニズムの課題を示している。また、インドネシアとフィリピンのVNRでは、「誰一人取り残さない」という2030アジェンダの理念に基づく指標モニタリングに関する具体的課題が示されており、特にフィリピンVNRには詳細な記述がみられた。しかし、これら2か国を除くと、フォローアップ・レビューに関する記述は概ね限定的であり、特にタイは表3に示した項目すべてにおいて何も示されていなかった。

先に概観したハウゲ報告で指摘されたVNRにおいて評価機能が十分認知されていないとの認識は、VNRレビュー結果からも見て取れる。インドネシアの状況については、前段の石田報告にて述べたが、インドネシアVNRではデータ集積・集計において障害者、移民に関するデータ集計に課題があり、複数の統計情報を統合する難しさについて言及されている。また、SDGsを評価の「機会」として捉え得る点は、ハウゲ報告とヴィジャヤラガバン報告で指摘されたが、インドネシアVNRでは、SDGsの主要な貢献の一つは、計測可能な開発成果の指標を精緻化すること、そして開発の利益を確実にすることだとの見解が示されている。フィリピンVNRでは、自国のSDGsフォローアップ・レビューに関する課題について詳細に言及している。具体的には、ターゲット・指標の定義が不明瞭なものがあること、複数の目標の概念に重複や矛盾がみられ

ること、民間セクターを含む多様な情報源からのデータの調和化、指標（推計値）の算出手法、指標のモニタリングに必要なデータの生成・集計等が挙げられている。また、モニタリングのための予算が不足している点、技術支援が必要とされる点も併せて言及されている。

本節では、コロンボ会合における参加者のアンケート回答と各国のVNRから、2017年時点におけるアジア諸国におけるSDGsフォローアップ・レビューのメカニズムの現状と課題について検討を行った。その結果をまとめると、第一にアジア諸国におけるSDGsのモニタリングの枠組み形成は各国ですすめられており、その中では各国個別の状況を反映させる「ローカライズ」の要素も含まれている。第二にレビュー結果のフィードバック体制についての各国での取り組みはモニタリング体制と比して形成途上であるといえる。第三に各国でのフォローアップ・レビューの枠組み形成には進展がみられるが、指標データの集積・集計、新設指標のデータの取得手法等に課題を抱えており、フィリピンやモルディブの課題に示されているように、技術的専門性や、財源が不足している状況もある。

前段で述べたハウゲ報告を通じて、2030年アジェンダをSDGsレビュー枠組の強化や、国家評価能力の開発、革新的な技術の利用、情報へのアクセスの改善の「機会」として捉え得ると述べたが、アジア諸国を含む開発途上国にとって、SDGsの時代が真の「機会」となり得るか否かは、いかに予算や人的資源の制約を取り払っていけるかに依るだろう。そして予算制約や人的資源の制約を取り払うためには、パートナー（ドナー）の支援は欠かせないものだといえる。

また域内間での南南協力、ドナーを加えた三角協力¹⁹を通じた相互学習の機会を通じて、先進的な取り組みから得られた学びの機会を設けていくことも重要である。このような機会を増やしていくためには、国際機関等の支援や、我が国ODAを活用した南南協力や三角協力を促進する取り組みも期待されるところである。

6. おわりに

複雑さが新たな日常になった（Complexity is new normal）現代社会において採択された2030アジェンダが目指す「誰一人取り残さない」世界を実現すべく掲げられたSDGsは既述の通り複雑である。

本稿ではSDGsの時代における開発途上国の評価能力構築について、国際的潮流、開発途上国の評価能力構築に向けた我が国の取り組みを通じて明らかになった開発のパートナーである開発途上国側の評価能力についての現状と課題を概観し、2030アジェンダとSDGsの時代における評価の「課題」と「機会」について検討を加え論じた。

SDGsのモニタリング・評価に際しては、指標の曖昧さ、統計基盤の不備、測定の難しさが指摘されており、この状況を改善するためには各国の事情に即した指標の設定、データ収集・集積、新たな評価手法の導入等の多くの「課題」を開発途上国に突きつけている。しかしながらこの「課題」は、見方を変えれば評価能力を構築する「機会」でもある。国際的評価コミュニティはこの「機会」により良く学習し、付加価値を高めることが求められている。

開発途上国の評価能力構築の重要性はパリ宣言、それに続く2030アジェンダの採択、SDGsの合意以降より一層国際社会で認知されることとなった。また、2030アジェンダの「誰一人取り残さない」理念は途上国側の評価能力構築に加え、参加型評価等の新しい評価手法導入の必要性を喚起する。

開発途上国の評価能力構築に向けた我が国の取り組みは、現在では外務省ODA評価室の「ODA評価ワークショップ」や「被援助国政府・機関等による評価」を通じた支援が中心であるが、過去にはJICAの技術協力も行われている²⁰。また近年では、石田報告にあるようにJICAの調査スキームを活用した支援が実施されており、新しい取り組みが開始されている。しかし、本稿5. で示したように、開発途上国側で必要とされる支援のニーズは膨大であり、今後は本格的な技術協力スキームを通じた評価制度の構築や

人材育成支援が期待されるところであり、そこでは南南協力や三角協力への支援も有効なツールとなろう。2030アジェンダの「フォローアップとレビュー」においても、「地域レベルでのフォローアップ・レビューは自発的なレビューを含む相互の学び、共通のターゲットに関する成功例と議論を共有する有益な機会となり得る」（国際連合2015、p.34）と示されており、上述の南南協力や三角協力は地域レベルでの相互の学び・機会の提供に貢献するといえる。

各国行政機関へのODAを通じた制度構築・人材育成支援に加え、学术界や評価協会を含む評価コミュニティへの支援も不可欠であろう。各国の評価能力の底上げには評価専門家個人の能力開発もまた必須であり、この点において評価コミュニティの果たす役割は大きい。とりわけ、最終受益者を視野に入れた「誰一人取り残さない」ための評価手法の研究・開発と開発途上国

での同評価の実践、そこからの学びを共有する相互学習の機会の提供・拡大への貢献が期待される。学習成果が国際的な評価フォーラムにて発信されることにより、SDGs達成に向けた国際的な議論への貢献も期待できよう。

最後にSDGsのユニークさは、先進国を含む全ての国が対象となっているその普遍性にある。開発途上国での経験は我が国を含む先進国のSDGs評価にも活用可能であり、また逆も真なりである。この点において、我が国におけるSDGs達成と、我が国のODAを通じた評価能力構築の間の橋渡し—その一貫性と整合性の確保—への貢献が期待されるのは、日本評価学会に代表される評価コミュニティであろう。評価コミュニティでの議論を通じ蓄積されるSDGs評価の理論と実践に関する知見は、SDGsとその評価の時代における開発途上国の評価能力構築へ大きく寄与すると考えられる。

表2 アジア大洋州諸国におけるSDGsのフォローアップ・レビューメカニズム（2018年）

	タイ	ネパール	パキスタン	バングラデシュ	フィリピン	ブータン	モルディブ	モンゴル
フォローアップ・レビューの枠組み								
SDGs 達成に向けてのロードマップが存在するか	○	◎	◎	◎	—	○	○	○
フォローアップ・レビューのデザイン・仕組み								
SDGs と国家開発政策成果との関連づけがなされているか	○	◎	◎	◎	◎	×	○	○
指標に必要なデータ収集方法が確立しているか	○	◎	◎	◎	○	○	◎	○
データ集積・集計システムが確立しているか	○	◎	◎	◎	○	○	○	○
レビュー結果のフィードバックの仕組み								
レビュー結果の各省庁へのフィードバックの仕組みが確立しているか	×	×	×	◎	○	○	×	×
国民へのフィードバック体制が確立しているか	×	×	×	◎	○	○	×	×
関係機関の巻き込み								
各省庁での SDGs 実現に向けてのアクションプランが策定されているか	○	○	○	○	×	×	×	○

(◎：取り組みが完了している、○取り組みが開始されている、×取り組みは開始されていない)

(出所) 筆者（日野）作成。

表3 自発的国別レビュー（VNR）から見る
各国のSDGsのフォローアップ・レビューメカニズムの現状と課題

国	指標の設定	データの集積・集計	国家開発政策への統合	SDGを「機会」と捉える	課題
インド	・フラッグシッププログラム、目標、ターゲットの関連省庁への割り振りを実施 ・州政府においてもSDGs指標のためのロードマップが策定をすすめている	記載なし	記載なし	記載なし	記載なし
インドネシア	・87 グローバル指標と234の代替指標を設定	・障害者、移民等のデータ集計には課題がある（複数の統計情報を統合する必要がある）	・今後策定される国家中長期計画にSDGsは組み込まれる予定	・SDGsの主要な貢献の一つは、計測可能な開発成果の指標を精緻化すること、そして開発の利益を確実にすること	記載なし
タイ	記載なし	記載なし	記載なし	記載なし	記載なし
ネパール	記載なし	記載なし	記載なし	記載なし	・脆弱なデータベース、性別、年齢、社会集団別データが存在しない
バングラデシュ	・70のグローバル指標のデータがあり、108の指標については十分ではないがデータがある ・残り63指標については、メカニズムを新たに構築する必要がある	記載なし	記載なし	現在 M&E システムを構築	・SDGsのローカライズ ・データの入手可能性、データ管理
フィリピン	・2016年6月までに8回のワークショップを実施 ・241の指標と自国の状況を勘案し74の追加指標を設定 【課題】 ・241の指標のうち121の指標については、定期的に把握されていない	【課題】 ・民間セクターを含む多様な情報源からのデータの調和化 ・推計値の算出手法 ・収集されているが生成されていないデータの存在 ・新たな集計作業が必要な指標データが存在する（雇用、ディーセントワーク、税、等） ・地方部からのデータの集積のメカニズムがない	・SDGsに関する同国のプライオリティは国家政策（中長期政策を含む）に基づいている	記載なし	・特定や測定が困難なターゲット・指標の存在（ターゲット、指標の定義や分類が不明瞭なものがある。また複数の目標の概念の重複や矛盾がみられる）。 ・多様な関係者（政府機関を含む）の調整 ・モニタリングのための資金が不足している（予算制約） ・指標を捕捉し、SDGsをモニタリングしていくためには、国内及び国外からの技術支援が必要とされる

マレーシア	・指標測定に必要なデータギャップ分析を実施	記載なし	記載なし	記載なし	・データの入手可能性の強化、データギャップの補填
モルディブ	・国家統計局によるデータの確認作業が終了 ・指標のプライオリティ付けを実施中	記載なし	記載なし	記載なし	・技術的専門性の不足、行政組織間の連携強化

(出所) 各国VNRをもとに筆者（日野）作成。

【課題】には、「指標の設定」及び「データの集積・集計」に関するセクションにおいて「課題」として示されていた事項を示した。最右列には、「SDGsのフォローアップ・レビューメカニズム」の課題として示されていた事項を示した。

注記

- 1 HLPFは、国連経済社会理事会主催（閣僚級、毎年）と国連総会主催（首脳級、4年毎。今回は2019年）がある。
- 2 <https://www.oecd.org/dac/DAC-HLM-2017-Communique.pdf>を参照。
- 3 <http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm>を参照。新たな評価クライテリアは2019年12月のDAC定期会合で採択された。
- 4 2018年には47か国、2019年には42か国、2020年には47か国がVNRを実施している。
- 5 Third International Conference on National Evaluation Capacities 2013。
- 6 Forth International Conference on National Evaluation Capacities 2015。
- 7 大臣官房ODA評価室から筆者（村岡）が参加。
- 8 <http://www.oecd.org/dac/evaluation/19thnetworkmeetingdocuments.htm>を参照。
- 9 <http://www.oecd.org/dac/evaluation/20thnetworkmeetingdocuments.htm>を参照。
- 10 http://www.mofa.go.jp/mofaj/gaiko/oda/kaikaku/hyoka/siryo_3_b.htmlを参照。
- 11 http://www.mofa.go.jp/mofaj/gaiko/oda/about/hyouka/page23_000809.htmlを参照。
- 12 南南協力とは、開発途上国の中で、ある分野において開発のすすんだ国が、別の途上国の開発を支援すること。JICAでは、「開発途上国が相互の連携を深めながら、技術協力や経済協力を行いつつ、自立発展に向けて行う相互の協力」と定義している。
- 13 <http://www.mofa.go.jp/mofaj/gaiko/oda/about/hyouka/>

[page22_000236.html](http://www.mofa.go.jp/mofaj/gaiko/oda/about/hyouka/page22_000236.html)を参照。

- 14 http://www.mofa.go.jp/mofaj/gaiko/oda/about/hyouka/page22_000828.htmlを参照。
- 15 インド、インドネシア、ミャンマー、マレーシアを除く。
- 16 本アンケートでは、参加国のモニタリング及びレビューの仕組み作りの現状と課題について一問一答の形式で参加者が回答する形式を取った。
- 17 Voluntary National Review。
- 18 SDGsモニタリング・フォローアップを所掌する組織に属する参加者は、フィリピンのみ。部分的に所掌する組織に属する参加者は、バングラデシュ、モルディブ、モンゴル、ネパール、パキスタン、タイ。ブータンは、当該業務を所掌しない組織に属している。
- 19 三角協力とは、先進国や国際機関が、途上国が他の途上国に対して行う南南協力を資金・技術・運営方法等で支援することを指す。
- 20 JICAは2021年1月から遠隔オンデマンド課題別研修「質の高い事業評価デザイン、実施及び制度構築のための能力強化」コースを開始した。

参考文献

- 外務省（2007）「第二章 国際的な援助潮流と日本の取組」、外務省『ODA白書 2007年版』、27-44
- 外務省（2016a）『ODA評価年次報告書 2016』
- 外務省（2016b）『フィリピンの防災分野における日本のODA評価＜概要＞』、
<http://www.mofa.go.jp/mofaj/gaiko/oda/files/000178990.pdf>
- 外務省（2017）『ODA評価年次報告書 2017』
- 国際連合（2015）『我々の世界を変革する：持続可能な

- 開発のための2030アジェンダ（外務省仮訳）
<http://www.mofa.go.jp/mofaj/files/000101402.pdf>
- Economic Planning Unit, Malaysia (2017) . *Sustainable Development Goals Voluntary National Review 2017*.
<https://sustainabledevelopment.un.org/content/documents/15881Malaysia.pdf>
 (2018年1月16日アクセス)
- Government of Nepal (2017) . *National Review of Sustainable Development Goals*.
<https://sustainabledevelopment.un.org/content/documents/15881Malaysia.pdf>
 (2018年1月19日アクセス)
- Government of the People's Republic of Bangladesh (2017) . *Eradicating poverty and promoting prosperity in a changing world Voluntary National Review (VNR) , 2017*.
<https://sustainabledevelopment.un.org/content/documents/15826Bangladesh.pdf>
 (2018年1月19日アクセス)
- Republic of Indonesia (2017) . *VOLUNTARY NATIONAL REVIEW (VNR) - Eradicating Poverty and Promoting Prosperity in a Changing World*.
<https://sustainabledevelopment.un.org/content/documents/15705Indonesia.pdf>
 (2018年1月19日アクセス)
- Republic of Maldives (2017) . *Voluntary National Review for the High Level Political Forum on Sustainable Development 2017*.
<https://sustainabledevelopment.un.org/content/documents/15891Maldives.pdf>
 (2018年1月19日アクセス)
- Republic of the Philippines (2016) . *Voluntary National Review at the 2016 High-Level Political Forum On the Sustainable Development Goals (SDGs)* .
https://sustainabledevelopment.un.org/content/documents/10765PH%20VoluntaryNationalReview_FINAL.pdf
 (2018年1月19日アクセス)
- The Government of India (2017) . *On the Implementation of Sustainable Development Goals*.
<https://sustainabledevelopment.un.org/content/documents/16693India.pdf>
 (2018年1月19日アクセス)
- United Nations (2015) . *Promoting inclusive and accountable public services for sustainable development. (A/RES/69/327)* .
<https://undocs.org/en/A/RES/69/327>
 (2018年1月19日アクセス)

(2021.7.6 受理)

Evaluation Capacity Development in SDGs Era

Keiichi Muraoka

Director, ODA Evaluation Division,
Minister's Secretariat,
Ministry of Foreign Affairs of Japan
Muraoka.Keiichi2@jica.go.jp

Ruiko Hino

ODA Evaluation Division,
Minister's Secretariat,
Ministry of Foreign Affairs of Japan
hino.ruiko@gmail.com

Abstract

In this paper co-authors present the global trends on Evaluation Capacity Development (ECD), Japan's contribution to ECD through ODA, ECD in partner countries and their challenges. Based on these observations, they discuss ECD in SDGs era.

The 2030 Agenda for Sustainable Development with its "leave no one behind" motto and Sustainable Development Goals (SDGs) call for ECD in developing countries and new evaluation methodologies such as participatory evaluation. These are opportunities as well as challenges for everyone concerned in evaluation. The needs for ECD in developing countries are huge and in order for them to take advantage of these opportunities, the support for ECD from donors including Japan are indispensable.

Keywords

SDGs, Evaluation Capacity Development, Ministry of Foreign Affairs, ODA, DAC Network on Development Evaluation

【研究論文】

SDG 教育目標にみる理念志向ターゲットの評価に関する一考察：測定可能性（measurability）から評価可能性（evaluability）へ

米原 あき

東洋大学

yonehara@toyo.jp

要 約

SDG評価の特徴でありかつ困難のひとつは、抽象度が高く質的な、理念志向のターゲットが含まれていることである。本稿の目的は、SDG理念志向ターゲットの評価を捉える理論枠組みを提示することと、その枠組みに示される「ローカル - 形成評価」の実効性を検討することである。そのための方法論として、理論考察とアクション・リサーチを併用する。

従来の画一的な定量指標では捉え難い理念志向ターゲットの登場により、測定可能性の追究から評価可能性の検討へと発想をシフトする必要性が生じている。評価可能性を高めるためには、その国やその社会の文脈を考慮して、抽象的なグローバル目標を実践現場の活動へローカル化する手続きが必要になる。本稿では、SDG理念志向ターゲットの典型例であるSDG4.7のESDに注目し、公立小学校においてESD評価のアクション・リサーチを行った。その成果から、「ローカル - 形成評価」におけるセオリー評価の意義と役割が明らかになり、中でも協働型のマネジメントが実効性の鍵を握ることが分かった。

キーワード

SDG4.7、評価可能性、プログラム評価、セオリー評価、ESD

1. はじめに

(1) 問題の所在

2015年9月の国連サミットで持続可能な開発目標（Sustainable Development Goals: SDG）が採択されてから5年が過ぎ、2030年までの進捗や成果をモニタリング・評価する方法についての議論と試行錯誤がますます活発になっている。教育目標に着目すれば、ミレニアム開発目標（Millennium Development Goals: MDG）時代には「初等教育の完全普及」などの具体的かつ測定可能な目標が掲げられていたのに対し、SDGの教

育目標には、持続可能なライフスタイルやグローバルな市民性など、従来のグローバル目標には直接的には取り上げられることのなかった抽象度の高い「理念」が取り込まれたことにより、その評価が著しく困難になっている（Yonehara et al. 2017）。

国連統計局（United Nations Statistics Division: UNSD）では、当初、SDGの評価指標を、既にデータの収集が進んでいるTier 1、合意された指標はあるがデータ収集が進んでいないTier 2、そして指標の開発が課題となっているTier 3に三分類して指標を整理しており、理念的な課題に関す

る指標の多くはTier 3に分類されていた（UNESCO-UIS 2016；UN IAEG-SDGs 2020）。その後、Tier3指標については、UNSDが中心となって再定義や整理が進められ、2020年7月の指標の見直しをもって、SDG評価指標のリストからTier3指標が無くなることになった（UN IAEG-SDGs 2020）¹。SDGの開始からこの5年ほどの間に、開発指標による達成度の測定という観点から、開発評価に関する議論が急がれてきたことが分かる。

SDGのような壮大なグローバル目標のモニタリング・評価を実行する際に、グローバルに共有可能な指標が必要不可欠であることは言うまでもない。一方で、従来の開発目標には無かった、理念志向の開発目標の評価について考える際にも、「測定可能な指標の開発」という従来のアプローチを同様に適用することは妥当なのだろうか。

後述するように、戦後70年近くに及ぶ国際開発の歴史の中で、これまでのグローバル目標の評価は、「どのような指標で測るか」という測定可能性（measurability）の観点から議論されることが主であった。戦後の国際開発ではGNPやGDPなどの経済指標が開発評価の中心的な役割を果たし、経済開発から人間開発へのシフトが叫ばれるようになった90年代以降も、後述する人間開発指数（Human Development Index: HDI）やジェンダー不平等指数（Gender Inequality Index: GII）などの複合指数が開発される方向で、すなわち、あくまでも測定可能性を探る方向で、評価の議論が進められてきた。この傾向はSDG時代を迎えた現在も大きくは変化していない。

一方で、SDGの17目標のもとに設定された169のターゲットの中には、後述するSDG4.7のように理念的な志向性が強いターゲットが含まれている。これらのターゲットが志向する理念には、その社会、その文化の文脈に基づく解釈と、その解釈に基づいて何をどのように評価するのか／すべきかといった評価可能性（evaluability）の問題が付随する。ロッシ（2010、p.128）は「評価可能性アセスメントの3つの活動」として、以下の3点を挙げている。

①注意深くプログラムモデルを記述し、ゴールと目標を定義する。

②そのモデルが十分に定義されているか、評価可能かをアセスメントする。

③利害関係者の評価に対する関心と評価所見の活用可能性を確認する。

すなわち、単に評価が可能かどうか（②）という技術的な問題だけではなく、そもそも何を目指してその取り組みを行うのか（①）、また、その評価結果は誰のために、どのように役立つのか（③）という点も吟味する必要がある。理念志向ターゲットの評価には、測定可能性に関する技術的な議論に加えて、評価可能性の観点からの議論が求められていると言えよう。

しかしながら、SDGの理念的なターゲットの評価に関する意義やその特性について、理論的な検討を深めた先行研究は管見の限り見当たらない。ジェンダーや衛生といった各論の立場からMDGからSDGへの変化を議論した先行研究はあるが（Death and Gabay 2015; Fukuda-Parr 2016; Roche et al. 2017; 湯浅ほか 2017など）、SDGの理念的な側面をどのように評価の枠組みに取り入れるのかについての議論は見られない。三浦ほか（2017）は、SDGにおける指標とモニタリング枠組みについての現状を、保健医療分野を例に分析しているが、結論で指摘されているのは、包括的な指標開発の必要性、すなわち測定可能性の問題である。

また、自発的国別レビュー（Voluntary National Reviews: VNR）を活用して各国政府に変革的開発を促すべきとする提言や（D'Errico 2018a, 2018b）、システム・アプローチを使ってSDGの包括性を捉えなおそうとする試みなどは（Lim, Jørgensen & Wyborn 2018）、測定可能性に留まらない新たな視点を提供し、SDG評価の視野を広げることに貢献しているが、ここでもSDGの理念的なターゲットをどのように扱うべきかという議論はなされていない。

（2）本研究の目的と構成

したがって本研究では、SDG理念志向ターゲットの評価について、その特性や意義を再検討するための理論枠組みを提示することを目的とする。本研究は、SDGの169ターゲットのうち、特定のターゲットを理念志向ターゲットとして

厳密に再定義しようとするものではなく、様々なターゲットの中で、理念志向性の強いターゲットの評価を考える際や、あるターゲットの理念的な側面を尊重した評価を行いたいときの指針を提供することを目指すものである。本稿では特にSDG教育目標（Goal4）のターゲット7（以下SDG4.7）に注目する。

SDG4.7に注目する理由は以下による。SDG開始当初に設定された約230の指標のうち、Tier 3指標の多くがSDG4.7に集中していたことから、当初よりこのターゲットの評価の困難性が指摘されていた（UNESCO-UIS 2016）。さらにSDG4.7の文言には「持続可能な開発のための教育（Education for Sustainable Development：以下ESD）」、「人権」、「ジェンダー平等」、「平和及び非暴力的文化の推進」など、教育目標というひとつの目標の範疇に留まらない、包括的な理念が多数含まれている。したがって、SDG4.7の評価について検討することにより、教育目標以外の領域における理念志向ターゲットの評価への応用可能性も示すことができる。

さらに本研究では、SDG4.7の教育実践としてのESDに注目し、小学校におけるアクション・リサーチを通じて、理念志向ターゲットの評価活動の実効性を検討する。具体的には、プログラム評価理論（ロッシ2010；源 2016）におけるセオリー評価の役割を、評価可能性の観点から再検討する。

本稿の構成は以下の通りである。まず、国際開発の評価において中心的な役割を果たしてきた開発評価指標の歴史を概観する（2節）。次に、「指標による測定≒評価」と考えられてきたと言っても過言ではない、戦後の国際開発の評価史の中で、SDG理念志向ターゲットが登場したことによって、測定可能性から評価可能性への転換が求められていることを指摘する（3節）。そして、SDG理念志向ターゲットを捉える理論枠組みを提案し、その位置づけを確認したうえで（4節）、小学校におけるプログラム評価の実践事例をもとに（5節）、理念志向ターゲットの評価におけるセオリー評価の機能について考察する（6節）。

2. 開発評価指標の歴史と発展

戦後、コロンボプランを皮切りに始まった国際開発の歴史の中で、開発評価の中心を担ってきたのは経済指標である。世界銀行が公刊している『世界開発報告』の初版（1978年版）において、貧困撲滅プログラムの推進にむけてモニタリング評価とそのためのデータベースづくりが重要であることが語られており（World Bank 1978：36-37）、報告書の巻末には各国のデータが掲載されている。ここで主たる開発指標とされたのは国民一人当たりGNPであった。後の報告書ではGNPに代わってGDPやGNIなどの指標が考案され普及していくが、いずれも経済力を開発評価の基準に置く考え方であることに変わりはない。

現在も経済指標が主要な開発評価指標として大きな影響力を持っていることに変わりはないが、経済開発を開発の目的とする開発観を、「開発をGNPの成長・個人所得の増加・工業化・技術の進歩・社会の近代化などと同一視する狭い開発論」（セン 2000：3）であると批判する人間開発論の登場によって、開発指標の考え方にも変化がもたらされた。

国連開発計画によって1990年から刊行が始まった『人間開発報告書（Human Development Report：HDI）』では、「人間開発指数（Human Development Index：HDI）」と呼ばれる複合指数が開発指標として採用されている。HDIは、3つのディメンション、すなわち、健康的な生活（Long and healthy life）、知識（Knowledge）、適正な生活水準（A decent standard of living）を考慮した複合指数で、それぞれのディメンションを代表する変数（出生時平均余命、平均就学年数、期待就学年数、一人当たりGNI [PPP US\$]）から算出された3つのインデックスの幾何平均として定義される（UNDP 2016：1-3）。HDIは、開発の評価を経済の側面からしか捉えることができない従来の経済指標と異なり、教育や保健を含むより包括的な開発指標として注目された。さらに、前述のジェンダー不平等指数（GII）や複合貧困指数（Multiple Poverty Index：MPI）など、様々な複合指数がHDIの影響を受けて考案された。

HDIやそれに続く多様な複合指数が、開発指標に広がりをもたらしたのは事実であり、それによって開発の目的を問い直すきっかけがもたらされた——すなわち、戦後の経済開発中心の国際開発から、人間開発へのシフトが図られた——という点は、複合指数による開発評価がもたらした重要なインパクトである。例えばJolly (2002: 266) は、HDIの登場をうけて、「統計学者たちは、統計学は「人権などに関する理念的な課題には（筆者注）貢献できないという誤解を捨てなければならぬ（筆者邦訳）」と訴え、統計指標が人間開発の推進に向けて人々を動機付ける働きを持つようになったことを歓迎している。

では、複合指数による評価というアプローチは、SDGの理念志向ターゲットの評価にも適合するのだろうか。複合指数は単一指標よりも高い包括性を実現しているとはいえ、国際機関などの評価専門家が策定した普遍的な指標を用いて、国際的に比較可能な相対評価を行うという基本的な構造において、従来の開発評価指標と本質的な違いはない（米原 2013）。さらに、異なる指標をウエイト付けして加算しても現実の複雑な様相は見えてこないという批判もある（ステイグリッツ2020）。SDGの理念志向ターゲットには、現実の様々な文脈に即した質的な要素が多分に含まれており、ユニバーサルな比較の視点では捉えられない要素をどのように評価のプロセスに取り込んでいくのかが問われている。次節ではこの点を評価可能性の観点から検討する。

3. SDG理念志向ターゲットの評価：測定可能性から評価可能性へ

(1) MDGからSDGへの変化の本質

2015年を目標年度に策定されたMDGは、8つの目標と21のターゲット、そして60の指標で構成されていた。これに対してSDGは17の目標と169のターゲット、そして231の指標へと大幅に拡大されたが（2020年7月現在：UN IAEG-SDGs 2020）²、SDGにおける開発目標の拡大は量的な意味に留まらない。教育目標であるMDG2及びSDG4（タ

ーゲット7）に注目すると、表1に示すように、その本質的な違いが浮き彫りになる。

MDG2の目標は「教育の普及」という量的な目標であったため、目標の記述内容と指標の定義が直接的に繋がっていた。事実、MDG2に連なる3つの指標は、表1のとおり、いずれも定量指標が設定されていた。これに対してSDG4が目指すのは、「包括的かつ公正な質の高い教育の確保と生涯学習の機会促進」という抽象度の高いもので、「質」への言及が明確になされている。さらに、特に抽象度が高く評価が困難と言われているSDG4.7が目指すのは、「持続可能な開発を促進するために必要な知識及び技能を修得できるようにする」ことであるが、ここに言われる「知識及び技能」が何を意味するのかは、それぞれの社会的文脈に大きく依存し、目標の記述には明示されていない。このような「知識及び技能」を育成する教育の方法として、表1の下線に挙げられているようなアプローチが提案されているが、例えば、「持続可能な開発のための教育（ESD）」や「人権、ジェンダー平等、文化多様性を促進するような教育」がどのような教育なのかについても、その国や地域の文化や宗教の違いを反映した、多様な解釈があり得るだろう。

さらに、MDG2が想定する教育の「場」が主として「学校」、特に初等教育学校であったのに対し、SDG4の指標には「すべての教育段階において」と明示されているうえ、ゴール記述には「生涯学習の推進」が謳われていることから、あらゆる教育機会がSDG4の「現場」となり得ることが分かる。学校教育の中でも特に義務教育段階に焦点を当てたMDG2では、初等教育や識字教育の社会的価値やその必要性について、一定の合意を想定することができ、また、この目標を達成するための主たる支援者が学校関係者であることも容易に想定できた。つまり、ゴール記述が直接的かつ明白に「指標」や「現場」や「ステイクホルダー」につながっていく構造を持っていた。一方、学校教育に限らないノンフォーマル教育など、あらゆる学びの機会を包含するSDG4は、多様なステイクホルダーがそれぞれに持つ、多様な価値のせめぎあいを内包する。したがって、そこに求められる「指標」も当然、

表1 MDG及びSDGにおける教育目標

	MDG2	SDG4 ^(注1)
ゴール	初等教育の完全普及の達成	すべての人に包摂的かつ公正な質の高い教育を確保し、生涯学習の機会を促進する
ターゲット	[2.A] 2015 年までに、全ての子どもが男女の区別なく初等教育の全課程を修了できるようにする (Goal2 のターゲットは A のみ)	[4.7] 2030 年までに、 <u>持続可能な開発のための教育及び持続可能なライフスタイル、人権、ジェンダー平等、平和及び非暴力的文化の推進、グローバル・シチズンシップ、文化多様性と文化の持続可能な開発への貢献の理解の教育</u> ^(注2) を通して、全ての学習者が、持続可能な開発を促進するために必要な知識及び技能を習得できるようにする
指標	・ 初等教育における純就学率 ・ 第 1 学年に就学した生徒のうち初等教育の最終学年まで到達する生徒の割合 ・ 15 ～ 24 歳の男女の識字率	[4.7.1] ジェンダー平等及び人権を含む、(i) 地球市民教育、及び (ii) 持続可能な開発のための教育が、(a) 各国の教育政策、(b) カリキュラム、(c) 教師の教育、及び (d) 児童・生徒・学生の達成度評価に関して、全ての教育段階において主流化されている程度

(注1) 紙幅の制約上省略するが、SDG4には7つのターゲット（4.1～4.7）と3つの手段ターゲット（4.a～4.c）が設定されており、MDG2.Aの内容はSDG4.1に含まれている。7つのターゲットと3つの手段ターゲットには合計12のグローバル指標と32の分野別指標が設定されており、MDG2の3指標と類似の指標はSDG4.1及び4.6に包含されている。SDG4.7のグローバル指標は上記の4.7.1のみで、他に5つのテーマ別指標が定義されている。

(注2) 下線部は筆者による。

(出所) 外務省（2019）、総務省（2019）、UN IAEG-SDGs（2020）を参考に筆者作成

多様な価値を反映したものとなり得る。

これらの取り組みを具体化するためには、これらのゴールやターゲットが具体的、個別的に何を意味しているのかを、それぞれの文脈や状況に応じて再解釈する必要、すなわち、評価可能な目標として再設定する必要が生じる。また、その再解釈・再設定のプロセスに関わるステイクホルダーが多岐にわたるという点も看過できない。つまり、多様な価値観をもつ多様なステイクホルダーが価値の再解釈・再設定の議論に参加できる公共空間を保障する必要も同時に生じているのである。

UNESCO（2016）はこのような再解釈のプロセスを「開封（unpacking）」と呼び、それぞれの国や地域コミュニティのレベルで、SDG4の抽象的な理念を実行可能な形に開封するためのガイドラインを提供している。この開封という手続きが必要になったという点が、MDG以前には求められることのなかった、SDG理念志向ターゲットの最大の特徴であると言えるだろう。そして、この開封という手続き概念を評価論の文脈で解釈すると、評価可能性の考え方が適用できる。

(2)「開封」プロセスとしての評価可能性アセスメントとセオリー評価

1970年代から広まった評価可能性アセスメント（evaluability assessment）は、元々、形成評価よりも総括評価の精度を向上させるための方法として考案された（Smith 1989）。しかしながら、評価可能性アセスメントのプロセスには、形成評価や計画策定などにも役立つ、以下の3つの有用性があると指摘されている（Smith 1989：12）。

- ①総括評価のツールとして：インパクト評価や効率性評価の前段階として、そのプログラムのどの部分が評価可能なのか検討する。
- ②形成評価のツールとして：そのプログラムをより効果的なものにするために、何を修正する必要があるのか検討する。
- ③計画策定のツールとして：そのプログラムのゴールを定義し、そのゴールを達成するために必要な活動や、それらの活動を量質共に十分なかたちで行うために必要な資源について検討する。

とくに3点目の特徴については、事業前評価（proactive evaluation）としての機能も持ち得るとされており、本稿の冒頭で紹介したロッシ（2010）の指摘にも通ずる。ここで議論されるのは「ゴ

ールの定義」や「そのゴール達成のための活動」であり、「修正や改善の方向性」である。つまり、「どのような指標で測るか」といった測定可能性を議論する前に、「そもそも何を求めて何を測ろうとするのか、その取り組みのゴールは何なのか」という評価可能性に関する議論を行う段階を明確に位置づける必要があるということである。

このような議論に立脚しながら、より包括的に取り組み全体の評価活動を体系化したのがロッシ（2010）のプログラム評価理論である。プログラム評価の5段階³に照らせば、セオリー評価と呼ばれる段階が、評価可能性を検討する段階となる。セオリー評価とは「プログラムに内包される理論が、目的に対してもっともらしく組み立てられているかどうかを検証し、プログラムの設計や戦略がアウトカムを達成するために妥当であるのかを評価するもの」（源 2016：40）である。つまり、ゴールの定義とそのゴールを達成するための手段や方策に論理的な齟齬がないかどうかを確認することがセオリー評価の役割であり、いわば、そのプログラムの屋台骨を組み立てる作業がセオリー評価にあたる。それゆえ、「プログラムの概念化が不十分であるか誤っている場合、たとえ概念化したものを適切に操作的に定義したとしても、プログラムが成功する見込みはほとんどない」とロッシは指摘している（2010：155）。「概念化したものを操作的に定義する」とはまさに指標の策定を指すと考えられるが、プログラムの概念化・理論化が不十分な状態で操作的定義を行っても、そのプログラムはうまくいかないと、ロッシは警鐘を鳴らしているのである。

評価可能性アセスメント及びセオリー評価の考え方に照らせば、「理念志向ターゲットを開封する」ということは、すなわち、ゴールの定義を明確にし、そのゴールを達成するために必要な資源について検討し、さらに目的と手段の間に論理的な構造のある計画を立てる、ということの意味する。これが実現して初めて「概念化したものを操作的に定義する」、つまり、測定可能性を検討し、適切な評価指標を考案することの意義が生まれ、評価可能性の高い、実行性あ

るプログラムが実現されるのである。次節では、測定可能性から評価可能性へのシフトを視野に入れた理論枠組みについて検討する。

4. SDG理念志向ターゲットの位置づけ

（1）開発評価を捉える理論枠組み

上述のようなひろがりをもつSDG評価を再定位するために、評価の目的と適用レベルを2つの軸とする枠組みを提示する（図1）。

まず目的軸であるが、Scriven（1991）によれば、評価の目的には大別して形成評価と総括評価があり、前者は内部関係者が改善を図るために必要な情報を提供することが目的であるのに対し、後者は外部者に対する説明責任を果たすことを評価の主たる目的とする。この定義に則り「形成評価 - 総括評価」という評価目的を両極とする軸を設定する。

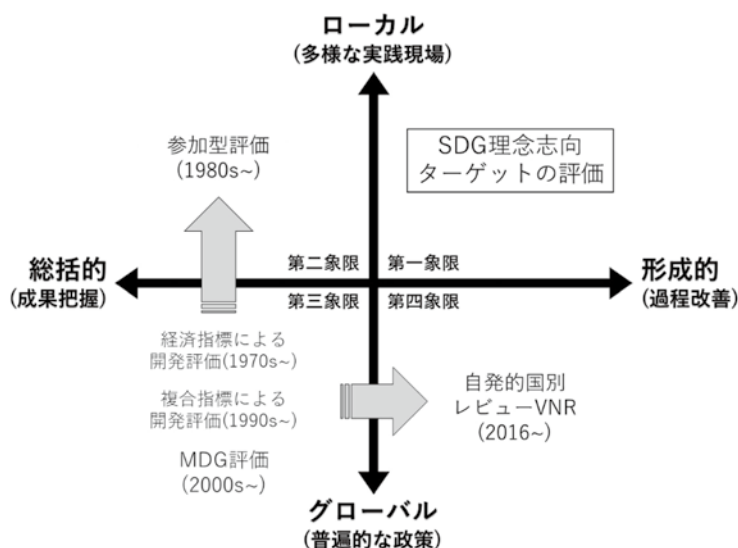
SDG評価の文脈でこの軸が示唆するのは、より一般的な比較可能性に対する期待や必要性の程度であると解釈できる。つまり、総括評価を目的とする立ち位置は、普遍性の高い指標で状況を測定して得られた、より科学的なエビデンスに基づいて成果を比較することを目的とする立場を示している。相対評価やランキングによる評価がここに当てはまる。

一方、形成評価を目的とする立場は、普遍的な指標による比較よりも、ある取り組み自体がもたらす個別の変化を確認し、改善に向けて具体的な課題を洗い出すことを目的とする立場を示している。したがって、この目的のもとに策定された指標は、必ずしも比較可能な一般性を備えている必要はなく、むしろ個々の文脈に応じた指標によって、絶対評価やレーティングによる評価が行われればよいことになる。

SDG理念志向ターゲットの評価は形成評価との親和性が高いと考えられる。つまり、普遍的画一的な基準では測り難い個々の文脈性や多様性、また、そこから生じる個別の価値観を反映した評価を重視する立場が、この極によって示される。

次に適用軸については、「ローカル - グロー

図1 SDG理念志向ターゲットの評価の位置づけ



（出所）米原（2016）より改変

バル」という軸を設定する。この軸は、評価を適用する際のレベルを示唆している。評価の諸活動は、特定のプロジェクトなど、具体的な現場で実践されることもあれば、SDGや各種の国際的な開発評価、あるいは国レベルの政策評価のように、ハイレベルの取り組みとして行われることもある。より現場に近いレベルで評価を実施する際には、現場の独自性や多様性に配慮した「開封」の手続きを重視する必要性が生じ、参加型の評価手法が適合しやすい傾向がある。一方、ハイレベルで評価を適用する場合には、個々の現場の独自性や多様性よりも、専門家による普遍的な説明力をもつ評価が求められる傾向がある。

適用軸に照らしてSDG理念志向ターゲットの評価を考えると、既述のとおり、この評価はローカルレベルで「開封」する必要がある評価活動であると言えるだろう。

（2）多様な開発評価の位置づけと展開

この枠組みに照らして開発評価の歴史を概観すると、戦後に始まった経済指標による開発評価は「グローバル - 総括評価」という第三象限の典型例であることが分かる。また、近年の国

際協力プロジェクトの評価については、相手国カウンターパートとの協働のもとにプロジェクトを進めていこうとする参加型のプロジェクト運営が主流になっていることから、第三象限からより「ローカル」へと適用レベルが移行しているといえよう。しかし同時に、プロジェクトの成果を、より科学的に頑強なエビデンスをもって示すことも求められる傾向が強まっていることから、近年の国際協力プロジェクトの評価活動は「ローカル - 総括評価」の第二象限を中心に展開していると考えられる。

もうひとつの最近の動向として、前述の自発的国別レビュー（VNR）が挙げられる。2016年より、国連経済社会局は、各国に自発的に自国のSDG進捗状況を報告することを推奨しており、VNRをデータベース化してオンライン上で公開することにより、SDG達成に向けてのピア・プレッシャーが生じることを期待している（UNDESA 2017）。VNRは国際的に共有される自己評価報告書の一種であり、各国の政府レベルで評価専門家によって作成される評価書のひとつであることから、その適用レベルは「グローバル」であると言える。同時に、自発的報告書であることから、その様式や評価基準などについては各国

の事情を尊重するという方針がとられており、国際比較などの相対評価は想定されていない。VNRでは、客観的な共通指標によって各国の取り組みを相対評価するのではなく、各国が自国社会の文脈に応じて、独自の改善に努めることが期待されている。したがってVNRのような評価活動は「グローバル - 形成評価」の第四象限に位置づけられる。

最後に、本稿の主題であるSDG理念志向ターゲットの評価は、「ローカル - 形成評価」として第一象限に位置づけられる。MDG評価は、8つの目標毎に設定された定量指標によって、グローバルレベルで成果を把握することを目的としており、必要な指標は国際機関が策定して収集していたという意味で、「グローバル - 総括評価」(第三象限)であった。これに対してSDG評価は、Tier1指標などによって従来の「グローバル - 総括評価」を維持しつつ、第一、第二、第四象限へと評価の守備範囲を大きく拡張している。中でも、上述の通り、SDG理念志向ターゲットの評価の特徴は「開封」のプロセスにあり、そこには多様なステイクホルダーが関与することになる。したがって、その評価活動は必然的にそれぞれの現場に根付いた文脈性の高いものとなり、同時に、多様な関係者を巻き込む参加型の評価となる。このような開封プロセスの登場によって、「ローカル - 形成評価」の第一象限が新たな開発評価の領域として拓かれた。

次節では、「ローカル - 形成評価」の事例として、筆者による、SDG4.7 (ESD) の評価をめぐるアクション・リサーチを手掛かりに、理念志向ターゲットの評価可能性について、実践の観点から考察する。

5. SDG理念志向ターゲットの評価実践：SDG4.7 (ESD) の評価とスクール・マネジメント

SDG理念志向ターゲットの評価は、「ローカル - 形成評価」であることが、上述の理論的検討から明らかになった。では、そのような評価は、現実の文脈においていかに実現しうるのだろうか。

か。本節では、SDG4.7のESDに注目し、「ローカル - 形成評価」の実効性を検討する。

また、ここで改めて表1のSDG4.7の指標に注目したい。指標4.7.1に示されているのは、「割合」や「数」ではなく、地球市民教育やESDが制度化されている「程度(原文extent)」である。つまりこれは、就学率や識字率のような、ある教育的介入の成果を問う指標ではなく、SDG4.7を実現するために必要な条件が整っているかどうかの程度を確認するサインポスト(道標)としての指標なのである。したがって以下の実践事例では、取り組みの「成果」を問うためではなく、「程度」を検討するための考察を試みる。

(1) ESD評価の課題

ESDは、2005年からの10年間、「国連ESD10年」のキャンペーンによって世界各地に普及した教育活動で、持続可能な社会の担い手の育成を目指している(北村、佐藤、佐藤 2019)。UNESCOによればESDによって育みたい能力として以下の6点が挙げられている(日本ユネスコ国内委員会 2021)。

- ・持続可能な開発に関する価値観(人間の尊重、多様性の尊重、環境の尊重など)
- ・体系的な思考力(問題や現象の背景の理解、多面的かつ総合的なものの見方)
- ・代替案の思考力(批判力)
- ・データや情報の分析能力
- ・コミュニケーション能力
- ・リーダーシップの向上

ESDは特定の教科ではなく、持続可能な社会を実現するために必要となる、上記のような汎用的能力を育成する教育活動の基本方針を示すものである。したがって、その具体的な形態は多岐にわたり(開発教育協会内ESD開発教育カリキュラム研究会 2010)、その多様性と柔軟性ゆえに評価が困難になっている(米原 2016)。つまり、ESDの理念を実現するための教育活動は、それぞれの学校や地域の環境の中で「開封」する必要があり、その成果を測定する以前の問題として、そもそも何をもってその学校のESDとするのか、その活動によって何を実現したいのかという評価可能性の問い直しが求められる。

ここでは、筆者が開校当初から支援している横浜市立みなとみらい本町小学校（以下、MMH小学校）で導入しているプログラム評価の実践に焦点を当てる⁴。

（2）みなとみらい本町小学校の概要

MMH小学校は、2018年4月に開校した公立小学校で、6学年15学級に410名強の子どもたちが学んでいる。開校宣言の中に「豊かな資源を活かし持続可能な社会の担い手を育む小学校として発展する」という文言が含まれている通り、開校当時から学校全体でESDに取り組んでいる。

しかしながら、開校当時はESDについての知見を持たない教員も多数おり、開校当初から小正和彦校長や外部の研究者などがESDについての研修を実施してきた。これらの研修と並行して行われてきたのが評価に関する研究会やワークショップである。MMH小学校の独自性は、ESDの考え方をスクール・マネジメントに取り入れ、全校的に活かそうとしている点にある。この点においてプログラム評価との親和性が高く、開校当初から積極的にプログラム評価が導入されてきた。以下、プログラム評価の流れに沿ってMMH小学校の取り組みを概観する。

（3）ニーズ評価・セオリー評価：「開封」プロセスとしてのロジックモデルづくり

MMH小学校では2018年4月の開校早々から重点研究としてESDが取り上げられてきた。具体的には、「持続可能な社会の担い手を育む小学校」として、どのような子どもたちを育てたいのか、その大きな目的（最終アウトカム）を意識化・言語化することから着手した。当事者であるMMH小学校の教員と、評価の伴走支援者である筆者の協働でワークショップを重ね、「グローバル人材を育てる」などのブラックボックス化した表現を次々と「開封」し、当事者である教員の言葉でロジックモデルに可視化していった。紙幅の制約上ここには掲載できないが、MMH小学校のロジックモデルはホームページで公開されている（みなとみらい本町小学校 2020a）。ここで重要なことは、ロジックモデルの詳細ではなく、それがMMH小学校のプログラム評価にお

いてどのような役割を果たしていたのかという点である。約3年の実践から、MMH小学校のESDロジックモデルには、以下の4点の機能があると言える。

① 上位アウトカムを言語化し自分事にする： 当事者意識の醸成

ESDの上位アウトカムは抽象的な概念で示されることが多い。例えば「グローバル人材を育成する」というアウトカムは、「グローバル人材とはどのような人物か」という定義が具体的に言語化されていなければ達成不可能かつ評価不可能な目的である。MMH小学校では、セオリー評価のワークショップの中で、「社会課題の解決に向けて実際に行動することができる子」、「社会に変化を起こすことができると信じられる子」の育成を教育目標にしようという方針が共有された。

毎回のワークショップでは個々の教員の具体的な取り組みや課題が話し合われ、それらの議論に基づいて上位の目的が設定された。上位アウトカムが、当事者である教員自身の言葉で具体化されたからこそ、ロジックモデルにある直接アウトカムや個々の活動が、日々の教育活動に紐づいた具体的なアイデアとして創出され、学校教育目標がお題目化することなく、日々の教育活動に活かされることとなる。また、その言語化の過程が関係者の間で共有されることに重要な意義があり、時間をかけてコミュニケーションすることによって納得感が生まれ、上位アウトカムがひとりひとりにとっての自分事となる。

② 上位アウトカムと具体的な活動との関係を 論理的につなぐ：取り組みのシナジー効果

ESDの教育活動は、その多様性ゆえに、個々の活動の質が高くても、学校全体・学年全体としてみたときに有機的な関連性に欠けるケースが多い。ロジックモデルで全体の関連性を可視化することで、個々の活動の位置づけや意味付けが明確になり、「様々な活動の単発的なやりっばなし状況」を「異なる活動間のシナジー効果の創出」へ転換することができるようになる。さ

らに、例えば「学年別の総合学習」と「遠足や運動会などの学校行事」など、種類の異なる教育活動と同じ目的に向かう手段としてみることができるようになり、異なる手段を連携させるなど、より有機的に教育活動を計画できるようになっている（みなとみらい本町小学校 2020b）。

③ 外部の関係者とのコミュニケーション・ツールとなる：外部関係者の参加促進

ESDの教育活動は、その包括性ゆえに、学校内外の多くの主体を巻き込むことがある。例えば、PTA活動や地域との協働に関する取り組みでは、PTAや地域の関係者とのコミュニケーションが必要となる。ここで求められるコミュニケーションとは、単なる事務的な打ち合わせではなく、どのような教育意図をもってこの活動を行おうとしているのか、この教育活動の成果をどのような目的につないでいきたいと考えているのか、といった活動の本質についての相互理解を深めるコミュニケーションである。しかしながら現実には、限られた時間の中でそのような対話を展開するのは難しい。

MMH小学校では、ESDロジックモデルをホームページで公開したり、後述する『ESD BOOK』（みなとみらい本町小学校 2020b）として保護者や地域の関係者に配布したりすることで、学校としての取り組みの共有を図っている。ロジックモデルを通じて視覚的・直感的に取り組みの全貌を示すことができ、2020年度の学校運営協議会では、PTAや地域の関係者から「自分たちに期待されていることが分かった」といった声や「このような活動があるならば具体的に○○といった機会の提供ができる」といった提案が挙がっていた。さらに、教育委員会や他校など、外部とのコミュニケーションにおいても、MMH小学校のESDの全体像を直感的に伝えることに役立っている。

④ 指標策定の基礎となる：調査データに基づく改善

ESD活動の評価に際しては、学校の教員がロジックモデルに照らして評価指標を策定し、それぞれの教育活動の成果をモニタリングしている。

上述の『ESD BOOK』はいわばモニタリング報告書のような役割も担っている。ESDの教育活動は個別性が高く多様であることから、その文脈・その目的に応じた指標を独自に作成する必要があるが、ロジックモデルは指標策定の際の道標となる。

ESD評価指標を作成する際に配慮すべき点は、既存の指標に捉われず、自分たちが実現したい価値をどう具体化するかについて、関係者の間で十分に議論を重ねることである。MMH小学校ではセオリー評価の一環として指標の検討を行った。このときにロジックモデルのすべての項目に指標を付けようとするのではなく、レバレッジ・ポイントとなりそうなところや、相乗効果が期待できそうなところ、あるいは取り組みのコアとして大切にしたいところに焦点を絞って検討した。

さらに、「指標」は必ずしも数値である必要はないとし、記述的な指標や、主観的な指標も積極的に取り入れた。その一方で、いわゆるテストスコアのような指標を排除する必要もなく、ロジックモデルに示された目的に照らして適切と判断されれば、例えば、ある教科のテストの点数が知識定着を測る指標のひとつとなることもありうる。「指標が多様である」という前提自体が、子どもたちや関係者に対して、「価値の物差しは多様でよい」というメッセージを発信することになる。MMH小学校では、ロジックモデルをもとに様々な指標が検討され、それらをアンケート調査票に反映して、次のような評価活動を行っている。

(4) プロセス評価・アウトカム評価：独自の指標に基づくデータ収集と分析

MMH小学校では、アンケート調査票を用いてデータの収集を行っている。そのデータを統計的に、あるいはテキストマイニングによって分析し、エビデンスに基づくプログラム改善を実践している。上述の通り、アンケート調査票に掲載されている質問項目は、ロジックモデルに基づいて策定された指標を文章化したものであり、したがって、収集されたデータは、ロジックモデルに示されたプログラムの進捗状況をモ

ニタリングするための情報源となる。定期的に同一のアンケート調査を行うことにより、期待する変化が実際に起こっているかどうかを、統計的な検定のもとに確認することもでき、そこから、具体的な改善に資する情報提供が可能になる。2019年度の調査は以下のとおり実施された（米原 2020）。

	第1回目（事前）調査	第2回目（事後）調査
時期	2019年10月23日	2020年1月15日
対象	低学年132人、 中107人、高82人	低学年131人、 中105人、高81人
設問数	低学年全7問、中・高学年全9問	

第一回目調査データと第二回目調査データのうち、数値的なデータには統計分析を適用し、自由記述欄のテキストデータにはテキストマイニングを適用した。紙幅の制約で一部に限られるが、以下に統計分析とテキスト分析の結果を例示する。なお、本節の目的は、これらの分析結果を詳述することではなく、MMH小学校で実施されているプログラム評価の流れを概観することである。したがって、ここでは分析結果の詳細には立ち入らないものとする。

① 数値データの分析例

ESDロジックモデルの活動のひとつとして「自分たちの考えや課題を更新しながら、発展的な課題解決学習を進めている」という項目があり、この目的を目指した様々な課題解決学習が行われた。その成果について、高学年を対象に「新しい課題や方法にもチャレンジしましたか（回答選択肢：とてもそう思う・そう思う・どちらとも言えない・そう思わない・まったくそう思わない）」という5件法による問いで確認したところ、事後の方が統計的に有意に向上していることが分かった（ $t=3.04$, $p<0.01$ ）。このように、ロジックモデルに対応した質問項目を設定し、各質問項目が対応するアウトカムの評価情報となるよう、アンケート票が設計されている。

② テキストデータの分析例

上述の課題解決学習については、「課題解決学

習を振り返って、自分にどんな変化があったと思いますか」という自由記述回答も設定された。テキスト分析の結果、事後の動詞語彙が顕著に豊かになっていることから、児童の変化が読み取れる（表2 網掛）。このように量的データだけでは不十分だと思われる項目には、質的な情報を収集するための質問項目が併設され、活動の改善に活用しやすいよう工夫されている。

表2 課題解決学習に関する自己評価

単語（動詞）	事前（%）	事後（%）
調べる	58	42
できる	50	50
考える	39	61
思う	26	74
増える	19	81
分かる	10	90
知る	10	90
生み出せる	0	100
変える	0	100
心掛ける	0	100
受けとめる	0	100
集まる	0	100
話す	0	100
持てる	0	100
持つ	0	100

（出所）筆者作成

(5) 改善に向けてのアクション

上述のような教育活動に対するアンケート自体は、一般的に多くの学校現場で実施されているが、ここで特筆すべきは、これらのアンケート項目がロジックモデルに紐づいており、MMH小学校のESD活動改善のための情報源になっていることである。これが可能になるためには、このアンケート票及びそのベースとなるロジックモデルが、教員自らの発案によって作成されたものになっている必要がある。外部の評価専門家や上位の行政組織が「外から」あるいは「上から」提示したもので、分析の結果や提言が自分事になりにくく、改善のアクションにつながり難い。

上記の調査結果を受けて、MMH小学校では、改善に向けて2つの具体的なアクションがとられた。ひとつは『ESD BOOK 2019』（みなとみらい

本町小学校 2020b) の出版である。ロジックモデルの策定から、具体的な教育活動の実施、アンケート調査、そしてその結果を踏まえての来年度の教育計画の策定という、一連の活動を担ってきた教員自身が執筆・編集者となって、自らの活動を振り返る機会となっている。さらにこの冊子を学校運営協議会のメンバーや保護者、地域住民、市教育委員会、他校の関係者などに配布することによって、MMH小学校のESD活動に対する理解と関心を深めてもらい、協働関係を育てることを意図している。

ふたつ目は、教員間で振り返りを共有し、具体的に次年度の計画改善に活かすための研修会である。2020年度末に開催された当該研修会では、アンケート調査によって得られた評価情報から、教員の認識と児童の実感に乖離のある項目や、思ったような成果に至っていない活動があることが明らかになり、次年度に向けての対策が練られる流れとなった。また、ESDロジックモデルは固定的なものではなく、毎年改善されることが前提となっており、研修会の議論を受けてロジックモデルの改訂も進められた。

形成評価の主眼は、評価そのものではなく、その結果を次の改善に活かすことにある。換言すれば、次の改善に活きるような評価情報が収集できる評価活動を行わなければ、形成評価を行う意味は半減する。理論的にはかく言われながら、現実には評価結果を現場の改善に結びつけるのは容易ではない。なぜなら、それを行うのは評価者ではなく、現場の当事者だからである。当事者にとって有意義な形成評価が、参加型・協働型になる必然性はこの点にある。MMH小学校では、当事者である教員による手作りのESDロジックモデルをもとに活動と評価が連動したことで「ローカル・形成評価」が適切に機能したものと考えられる。

6. 考察：セオリー評価の機能と可能性

MMH小学校の事例から「ローカル・形成評価」の実効性について考察すると、次の3点が要点として抽出される。

一点目はセオリー評価の重要性である。ESDのような、その中身が分かりにくいSDG理念志向ターゲットを実践のレベルで「開封」するためにも、また、そのような理念的な課題を当事者にとっての自分事にするためにも、当事者間のコミュニケーションを尊重したセオリー評価の実践が不可欠である。特に本節(3)に列挙したロジックモデルの4つの機能を活かすことが、のちの段階にも重要な影響を及ぼすと言える。

二点目は参加と協働が不可欠であるという点である。参加型評価において当事者の参加が必要であるというのは言うまでもないが、実践現場において、それが可能な組織文化とガバナンスが整備されているか否かが、参加と協働の実質性に多大な影響を及ぼす。学校現場の場合、校長ほか管理職のリーダーシップ、それも教職員が自由闊達に議論や活動を展開できる環境を築くためのサーバント・リーダーシップが鍵を握っている。

三点目は評価をマネジメントの一環として捉える視点が現場の関係者間で共有されているという前提条件である。特に日本の学校文化において、「評価」という言葉は往々にして「監査」や「管理」に近いニュアンスで受け止められることが多く、教職員が楽しみながら関われる活動として理解されることは少ない。しかしながら、形成評価が形成評価としての本来的な機能、すなわち現場に改善をもたらすという機能を発揮するためには、評価を、孤立した活動ではなく、組織やプログラムのマネジメントを駆動するエンジンとして捉える視点の共有が不可欠である。そしてそのような視点の共有と定着も、一点目に挙げたセオリー評価の過程で起こり得る。

ESDなどのSDG理念志向ターゲットの評価を実施する際、「開封」された理念の具体的な活動形態は、それぞれの現場の文脈によって大きく異なる。そのような多様な活動に対して「グローバル・総括評価」のアプローチを適用しても機能し難い。また、教育目的やその教育プログラム全体の整合性を検討せずに新たな評価手法や評価指標を導入しても、現場に混乱をもたらすことになるだろう。そこで、それらの教育活動

の評価可能性を高めるために、「注意深くプログラムモデルを記述し、ゴールと目標を定義する」（ロッシ 2010：128）ためのコミュニケーションが必要となる。

重要な点は、そのような本質的なコミュニケーションの機会が、活用可能な状態で実践現場に開かれていることである。セオリー評価としてスクール・マネジメントの一環にこのプロセスを位置づけることで、その活動の価値や目的について議論する公共のアリーナが保障され、そこから具体的な活動が醸成されていく可能性が担保される。従来は求められることのなかったセオリー評価の手続きが、SDGやESDの広まりと共に、様々な教育現場で求められるようになっていく。喫緊の課題として、このような評価活動を伴走支援する評価専門家の派遣制度など、実践現場を支える体制づくりが求められていると言えよう。

7. おわりに

本稿では、SDG理念志向ターゲットの評価を捉える理論枠組みを提示すること、そしてそこに示される「ローカル - 形成評価」の実効性を検討することを目的として考察を進めた。SDG評価の最大の特徴は、目標や指標の量的・質的な拡大ではなく、測定可能性の追究から評価可能性の検討へと発想をシフトする必要性が生じているという点にある。評価可能性を検討するためには、その国やその社会の文脈を考慮して、抽象的なグローバル目標を「開封」する手続きが必要になる。これらの特徴を踏まえ、SDG評価の枠組みを、目的軸と適用軸から構成される四象限で捉え直すと、SDG理念志向ターゲットの評価は、「ローカル - 形成評価」として特徴づけられる第一象限に位置することが分かった。さらに第一象限の典型例であるSDG4.7の実践事例から、参加と協働、マネジメントの一環としての評価、そしてそれらに通底するセオリー評価の重要性が指摘された。

図1の枠組みが示されたことにより、MDGからSDGへの「拡大」の意味が、開発評価史の観点

から問い直され、理念志向ターゲットの評価に加えて、VNRや国際協力案件の評価の動向など、SDG評価に関わる諸側面の相対的な位置づけが明らかになった。また、ESD評価の事例を検討することにより、「ローカル - 形成評価」の実践において、セオリー評価が重要な役割を果たすこと、そして、セオリー評価に重点を置いた評価活動それ自体が「当事者による公共の議論の場」の創造に寄与することが指摘された。加えて、セオリー評価を経て設定された指標が、SDG理念志向ターゲットの測定可能性を高めることも期待される。

今後の課題として、グローバル、ローカルそれぞれのレベルに積み残しがある。まず、本稿で取り上げたMMH小学校の実践事例は、いわば「良い逸脱」（河村、Singhal 2011）の一事例であり、今後MMH小学校のような取り組みをどのようにして水平展開していくのか、また、学校組織の評価文化や、学校関係者の評価的思考といった、「ローカル - 形成評価」を導入するための前提条件をどのように整えるのかという点については、更なる検討を重ね、方法論として体系化していく必要がある。同時に、SDG評価のグローバルな動向、特に、各国政府に対するアンケート調査方式で収集されることになったSDG4.7の評価データが、今後どのように収集・分析され、どのようなかたちで公表・活用されるのか、そして、それがローカルな教育現場にどのような影響をもたらすのか——現場のセオリー評価を促進するのか阻害するのか——についても注視する必要がある。

謝辞

本稿の修正に際して、匿名の査読者より、本研究の本質と今後の方向性に関わる重要なご指摘を頂きました。ここに記して感謝申し上げます。

注記

1 Tier3指標の多くは、その定義が再考されてTier2指

- 標に分類されることになった。例えば表1のSDG4.7の指標4.7.1は、(a)～(d)に関する細目について、それぞれの実施程度を各国政府にアンケート調査で聞き取り、総合点を付けるという方法で合意された(UNSD 2021b)。
- 2 2020年7月に見直された新たなフレームワークによれば、全体数としては274の指標が列挙されているが、うち12の指標については2～3の異なるターゲット間で同じ指標が共有されているため、固有の指標数としては231となる(UNSD 2021a)。
 - 3 プログラム評価の5段階とは「ニーズ評価、セオリー評価、プロセス評価、アウトカム／インパクト評価、効率性評価」の5段階を指す(ロッシ 2010)。
 - 4 MMH小学校は、2019年より、文部科学省『SDGs達成の担い手育成(ESD)推進事業：教育(学習)効果の評価と普及』のモデル校となった。本稿の一部は、当該事業の報告書による(米原 2020)。

参考文献

- 開発教育協会内ESD開発教育カリキュラム研究会 (2010)『開発教育で実践するESDカリキュラム：地域を掘り下げ、世界とつながる学びのデザイン』、学文社
- 外務省 (2019)『ミレニアム開発目標MDGs』(<https://www.mofa.go.jp/mofaj/gaiko/oda/doukou/mdgs.html>, 2021年3月24日)
- 河村洋子・Singhal, A. (2011)「社会の中の『良い逸脱』：Positive Deviance」、『熊本大学政策研究』、3：35-45
- 北村友人、佐藤真久、佐藤学 (2019)『SDGs時代の教育：すべての人に質の高い学びの機会を』、学文社
- スティグリッツJ. (2020)「グローバル化する世界における経済学者の役割とは」、『経済セミナー』、712：8-18
- セン A. (2000)『自由と経済開発』、日本経済新聞社
- 総務省 (2019)『持続可能な開発目標SDGs：指標仮訳』(http://www.soumu.go.jp/toukei_toukatsu/index/kokusai/02toukatsu01_04000212.html, 2021年3月24日)
- 日本ユネスコ国内委員会 (2021)『ESDで目指すこと(2) 育みたい力』(<https://www.mext.go.jp/unesco/004/1339970.htm>, 2021年3月24日)
- 三浦宏子、下ヶ橋雅樹、富田奈穂子 (2017)「持続可能な開発目標(SDGs)における指標とモニタリング枠組み」、『保健医療科学』、66(4)：358-366
- みなとみらい本町小学校 (2020a)『ロジックモデルを用いたESDの可視化』(<https://www.edu.city.yokohama.lg.jp/school/es/minatomiraihoncho/index.cfm/1,0,73,211.html>, 2021年3月24日)
- みなとみらい本町小学校 (2020b)『ESD BOOK 2019：未来創造』(<https://www.edu.city.yokohama.lg.jp/school/es/minatomiraihoncho/index.cfm/1,2556,c,html/2556/20200514-134454.pdf>, 2021年3月24日)
- 源由理子 (2016)『参加型評価：改善と変革のための評価の実践』、晃洋書房
- 湯浅資之、金子恵、安齋寿美玲 (2017)「持続可能な開発のための2030アジェンダ(SDGs)：ユニバーサリティとコンバージェンスの視点による考察」、『国際保健医療』、32(1)：17-26
- 米原あき (2013)「人間開発指数再考：包括的な開発評価への試み」、『日本評価研究』、12(3)：91-105
- 米原あき (2016)「『学び』の一環としての『評価』：協働型で行うプログラム評価の可能性」、『日本/ユネスコパートナーシップ事業ESDの教育効果(評価)に関する調査研究報告書』、岡山大学、52-61
- 米原あき (2020)「協働型プログラム評価によるESDスクール・マネジメント実践に関する調査研究報告書」、『文部科学省SDGs達成の担い手育成(ESD)推進事業：教育(学習)効果の評価・普及報告書』、35-63
- ロッシ P. 他 (2010)『プログラム評価の理論と方法：システマティックな対人サービス・政策評価の実践ガイド』、日本評論社
- Death, C. and Gabay, C. (2015) . Doing biopolitics differently?: Radical potential in the post-2015 MDG and SDG debates. *Globalizations*, 12 (4) , 597-612.
- D' Errico, S. (2018a) . VNR reporting needs evaluation: A call for global guidance and national action. *Briefing*, January, International Institute for Environment and Development. Retrieved Dec. 29, 2020, from www.jstor.org/stable/resrep16704.
- D' Errico, S. (2018b) . Embedding Evaluation in National Plans and Policies to Foster Transformative Development. *Briefing*, May, International Institute for Environment and Development. Retrieved Dec. 29, 2020,

- from <http://www.jstor.org/stable/resrep16718>.
- Fukuda-Parr, S. (2016) . From the Millennium Development Goals to the Sustainable Development Goals: Shifts in purpose, concept, and politics of global goal setting for development. *Gender & Development*, 24 (1) , 43-52.
- Jolly, R. (2002) . Statisticians of the world unite: the human development challenge awaits. *Journal of Human Development*, 3 (2) , 263-272.
- Lim, M., Jørgensen, P., & Wyborn, C. (2018) . Reframing the sustainable development goals to achieve sustainable development in the Anthropocene: a systems approach. *Ecology and Society*, 23 (3) , 22.
- Roche, R., Bain, R., and Cumming, O. (2017) . A long way to go: Estimates of combined water, sanitation and hygiene coverage for 25 sub-Saharan African countries. *PLOS ONE*, 12 (2) , e0171783.
- Scriven, M. (1991) . *Evaluation thesaurus*. CA: Sage Publications.
- Smith, M. F. (1989) . *Evaluability assessment: A practical approach*. Springer Netherlands.
- UNDP. (2016) . *Human development report: Technical notes*. UNDP. Retrieved March 24, 2021, http://hdr.undp.org/sites/default/files/hdr2016_technical_notes.pdf.
- UNESCO. (2016) . *Unpacking sustainable development goal 4 education 2030 guide*. UNESCO.
- UNESCO and UNESCO Institute for Statistics (UIS) (2016) . *Laying the foundation to measure sustainable development goal 4*. Quebec: UNESCO-UIS.
- United Nations Economic and Social Council (UNESC) (2016) . *High-level political forum on sustainable development: Progress toward the Sustainable Development Goals*. Retrieved June 2, 2021, from <https://unstats.un.org/sdgs/files/report/2016/secretary-general-sdg-report-2016--en.pdf>.
- United Nations Department of Economic and Social Affairs (UNDESA) (2017) . *Sustainable development Knowledge platform: Voluntary national reviews database*. Retrieved March 24, 2021, from <https://sustainabledevelopment.un.org/vnrs/>.
- United Nations Inter-agency and Expert Group SDG Indicators (UN IAEG-SDGs) (2020) . *Tier Classification for Global SDG Indicators*. Retrieved March 24, 2021, from https://unstats.un.org/sdgs/files/Tier%20Classification%20of%20SDG%20Indicators_17%20July%202020_web.v3.pdf.
- United Nations Statistics Division (UNSD) (2021a) . *SDG Indicators: Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development*. Retrieved June 2, 2021, from <https://unstats.un.org/sdgs/indicators/indicators-list>.
- United Nations Statistics Division (UNSD) (2021b) . *SDG Indicators: Metadata repository*. Retrieved March 24, 2021, from <https://unstats.un.org/sdgs/metadata/>.
- World Bank (1978) . *World development report: Prospects for growth and alleviation of poverty*. World Bank.
- World Bank (1996) . *World development report: From plan to market*. World Bank.
- Yonehara, A., Saito, O., Hayashi, K., Nagao, M., et al. (2017) . The role of evaluation in achieving SDGs. *Sustainability Science*, 12 (6) , 969-973.

(2021.7.8 受理)

A Consideration on Evaluation for Conceptual Targets of SDG4: Shift from Measurability to Evaluability

Aki Yonehara

Toyo University
yonehara@toyo.jp

Abstract

One of the prominent difficulties of SDG evaluation is rooted in its highly-abstract and quality-based conceptual targets. The purpose of this paper is to provide a theoretical framework to capture the SDG conceptual targets in the context of evaluation for global cooperation and to examine feasibility of “local-formative evaluation,” which appears in the framework.

Some of the SDG conceptual targets are difficult to measure by existing evaluation indicators. Therefore, it is necessary to shift an emphasis of our discussion from measurability to evaluability. In order to increase evaluability, abstract global goals and targets need to be localized by considering its own social context. This study focuses on ESD of SDG4.7 as a typical representative of the SDG conceptual targets and employs an action research method to analyze a case of ESD evaluation at a public primary school. This action research reveals the significance of theory evaluation as “local-formative evaluation,” which promotes co-productive management at the practical level.

Keywords

SDG4.7, Evaluability, Program evaluation, Theory evaluation, ESD

【研究論文】

SDGs 第 7 目標の評価

林 薫

文教大学

hayashi@bunkyo.ac.jp

要 約

SDGs第7目標（以下SDGs7と略）は汚染の少ない現代的エネルギーの普及と利用の効率性を通じて持続性を高めることを目指している。人々が貧困から抜け出すためにはエネルギーの使用が不可欠である一方で、エネルギー供給は枯渇性資源に依存することがまだ多く、第7目標がSDGs全体の成否の鍵の一つを握っていることは言うまでもない。SDGs7は他の多くのSDGs目標の実現と関連性が強い一方で、目標横断的なセオリーをSDGsは示していない。そこで、SDGs7は他の目標あるいは持続性全体との関係でどのようなセオリーを描きうるのかを、本稿のリサーチ・クエスチョンとした。SDGs7が重要な指標の一つとするエネルギー・インテンシティー（Energy Intensity）は持続性に示唆を与える指標になりうるのか、また、目標の一つである再生可能エネルギーの導入はアウトカムとして人々の暮らしにどのような変化をもたらしているかについて検討を行うことを通じて、セオリーの構築を試みつつ、SDGs7の評価、さらにそれを通じたSDGs全体の評価の課題を明らかにしていきたい。

キーワード

エネルギー、目標間の関係、セオリー、エネルギー・インテンシティー、GDP

1. SDGs7の目標と指標

(1) SE4ALLからSDGs7へ

2010年に国連は「すべての人々のための持続可能なエネルギー Sustainable Energy for All (SE4ALL)」を立ち上げた。このイニシアティブには、①エネルギーアクセスへの向上（2030年までに現代的なエネルギーサービスをいきわたらせる）、②エネルギー効率の改善（エネルギー効率を2倍にする）、③再生可能エネルギーの活用（世界全体のエネルギーミックスの中で再生可能エネルギーのシェアを2倍にする）の3本の柱があった。同時に、国連総会は2012年を「す

べての人々のための持続可能なエネルギー年 International Year of Sustainable Energy」とすることを決議した。SE4ALLの諸目標や指標は持続可能な開発目標SDGs7に踏襲されている。

国連が公表している「事実と数字」¹⁾ではSDGs7に関して以下のような認識を示している。

①世界人口の13%は依然として現代的電力を利用できていない、②30億人が薪、石炭、木炭、または動物の排せつ物を調理や暖房に用いている、③エネルギーは気候変動を助長する最大の要素であり、全世界の温室効果ガス排出量の約60%を占めている、④世帯エネルギーとしての可燃燃料使用による屋内空気汚染により、2012年

には430万人が命を失っており、女性と女兒はその10人に6人を占めている、⑤2015年、最終エネルギー消費に再生可能エネルギーが占める割合は5%に達した。

(2) SDGs7と指標

上に記した課題認識からSDGs7は付表1のように目標を設定している。要約すると、①近代的エネルギーおよび再生可能エネルギーの普及、②効率の改善、③そのための技術開発と普及、国際協力の3点である。これらの目標に対応した指標は付表1の目標のそれぞれ右に示す。指標に基づいた国別、地域別、世界全体の進捗状況は通称「ダッシュボード」と言われる、Sustainable Development Report で逐次確認することが可能である²。

(3) 第7目標の持続性概念

持続可能な開発の概念は「環境と開発に関する世界委員会」（委員長：ブルントラント・ノルウェー首相（当時））が1987年に公表した報告書「Our Common Future³」にある「持続可能な開発＝将来の世代の欲求を満たしつつ、現在の世代の欲求も満足させるような開発」が出発点になっている。しかし、将来の世代というのは、どこまで先を考えなければならないかについて具体的に示されているわけではない。このことは石油、天然ガスなどの枯渇性資源の役割が依然として大きなエネルギー分野において困難な問題を提起する。枯渇性資源においては現在の世代の欲求の満足は将来の世代の欲求の満足とゼロサムで競合し、いずれかの時点で、それ以降の世代の欲求を満足させることはできなくなる。

石油に関して2019年末の確認開発埋蔵量は1,734兆バレルで、前年より2兆バレル減少している⁴。この確認開発埋蔵量（R: Reserve）を年間の生産量（P: Production）で除したR/Pは全世界平均約50で、これが現在の生産（＝消費）レベルを続けた場合の「石油が使える年数」である。石油の生産・消費は価格の動向に影響を受ける。価格が高くなれば消費が抑制されることによりPが低下し、より高コストの石油採掘でも採算が取れるようになることから、石油探査や油田開発も進み、Rが上昇する。価格の循環的なプロセ

スで、過去半世紀R/Pは30～50で推移している。天然ガスもR/Pは40前後で推移してきたが⁵、最近ではシェールガスなどの非在来型天然ガスの開発が進み、より長く持ちこたえるという見方も出てきている。しかし、近い将来R/Pが急速に低下する懸念は常に付きまとっている⁶。今後の石油価格の高騰の際に代替的な再生エネルギーへの円滑な移行が進むという保証はどこにもない。SDGsは2030年までの目標を設定しているが、枯渇性の資源に関しては、かなりの長期間にわたるエネルギーの資源量と使用量のシナリオを描かない限り、持続性を捉えた目標とすることはできない（林 2017a）（林 2017b）。この点が2030年の達成を第一義的に目指す他の目標と異なるところである。

一人あたり年間GNIが3,995ドルを下回る低所得国・低位中所得国に住む人口は、世界人口の46.8%に達しており（2019世銀データベース）、生活改善のためにこれからもエネルギー使用量の増大は不可避かつ不可欠である。エネルギーの消費効率の改善が需給の緩和と価格の安定化を通じて貧困層にもたらす意味は大きい。目標となるのは生活水準の向上である。SDGs全体におけるSDGs7の役割は、各目標の実現のために必要とされる単位あたりのエネルギーを効率化によって削減し、全体としてエネルギー消費全体を抑制し、可能であれば削減するということになるだろう⁷。

SDGs7は従って、他の目標とのかかわりあいが必要であり、次節で見るようにいくつかのトレードオフもある。その前にもう一つ、SDGs7の持つ問題を指摘しておきたい。それは原子力発電についての言及を避けていることである。もちろん、原子力発電は極めて論争的な事項であり、コンセンサスの上に成り立ったSDGsにおいて言及がなかった事情は理解できる。原子力発電所に関しては、事故発生時の被害の大きさはもちろんのこと、半減期が数万年に達する使用済み核燃料や放射性廃棄物の安全な処理方法が争点になっている。数万年先の「将来の世代」に配慮するのは正しい。それならば、化石燃料を含むエネルギー全体に関しても同じスパンで考えるべきではないか。SDGsに超長期の視点が必

要な所以である。

(4) SDGs7の他目標とのトレードオフ

SDGs7の評価の観点からは他目標とのトレードオフが問題になる。SDGsの諸目標の間には相乗効果をもたらしたり相互補完したりする理想的な関係だけではなく、競合、矛盾する関係があることも指摘されている。International Council for Science (2017) によるSDGs7に関する目標間の関係分析結果を付表2に示す。要点は以下の通りである。

貧困削減目標 (SDGs1) との関係では、SDGs7が目標とする近代的なエネルギーの普及は相互補完関係になる。また、再生可能エネルギーの普及が気候変動を緩和することができれば、貧困層にとっては変動の影響の緩和につながる。しかし、再生可能エネルギーの半ば強制的な普及がエネルギー価格の高騰をもたらすとすれば、貧困削減目標とは矛盾する。例えば、太陽光発電について、高めに設定されたFIT (Feed in Tariff) 価格が、太陽光発電設備を導入できる比較的富裕な層に対して、それができない低所得層が割増された電気料金の支払いを通じて補助を行うという「逆進性」が指摘されていた (朝野 2015)。

飢餓・食糧 (SDGs2) に関しては、女性、先住民、家族経営の農林水産業などエネルギー集約的でない伝統的生産様式がSDGs7と相互補完的であると評価される一方で、バイオ燃料生産と食料生産は競合するとされる。化石燃料の使用削減は、採掘に利用する水の削減を通じて、水の利用可能性拡大 (SDGs6) と相互補完的であり、海水淡水化に際しては、エネルギー非集約的なフィルターなどの重点的な技術開発と使用拡大でSDGs7と両立する。

雇用・経済成長目標 (SDGs8) との関係では成長とエネルギー消費のデカップリングが課題である。エネルギーの普及は、学習や労働時間にポジティブな影響を与える。この点については第3節で紹介するミャンマーの調査からも明らかである。エネルギーの使用削減はエネルギー産業での雇用喪失をもたらす恐れがあり、これは2020年の米国大統領選挙でも争点の一つとなった。炭素 (炭化水素) の燃焼は必然的に二酸化

炭素を発生させるので、気候変動目標 (SDGs13) とは入口、出口の関係に立つ。このほか、上記の表には掲げていないが、責任ある生産、消費 (SDGs12) も表裏一帯の関係に立っている。

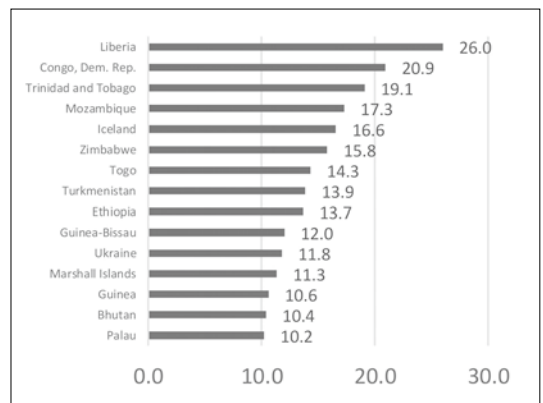
このように、SDGs7と他の目標は一致したり補完したりするものが多い一方、矛盾、相克するものもある。SDGs7の評価は他の目標とセットで行う必要がある。

2. エネルギー・インテンシティー (Energy Intensity) とSDGs7の評価

(1) エネルギー・インテンシティー

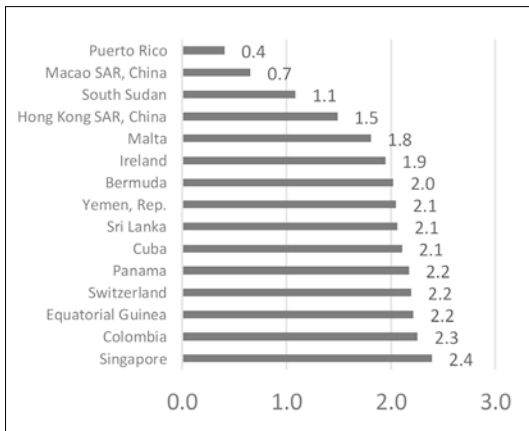
SDGs7が重要な指標として用いるのがEnergy Intensity (以下EIと略) である。EIは購買力平価で換算されたGDP1単位あたり使用されたエネルギーの熱量で示される (単位はメガジュール MJ)。SE4ALLにおいてエネルギー効率の指標として使われており、SDGsもこれを踏襲している。EIは、経済活動の「燃費」ということができる。2015年のEIの上位15か国と下位15か国をそれぞれ図1、図2に示す。

図1 EI高数値 (GDPに比較してエネルギー消費が多い) 上位15か国



(出所) 世銀データベースより筆者作成 (2021年3月21日)

図2 EI低数値（GDP に比較してエネルギー消費が少ない）下位15カ国



（出所）世銀データベースより筆者作成（2021年3月21日）

データを見る限りでは地域別、所得階層別の特徴は明確ではない。EIに影響を与える要因は複雑である。ただ、先行研究のEIの分析からエネルギーに関するさまざまな政策含意を得ることはできる。

Filipovic, et al. (2015) はEU28カ国の1990年から2012年までのパネルデータを用い分析し、EIを改善させる要因としてエネルギー価格、1人当たりのGDPなどを、悪化させる要因としては国内エネルギー消費および1人当たりのエネルギー使用量の増加などを挙げる。特に重要なものは電気料金であり、デンマーク、ドイツ、イタリアのEIが低いのは、電気料金が低いことによるものであるとする。エネルギー消費に対する補助金の削減は既にSDGs7の目標の一つとして組み込まれている。また一般的なエネルギー補助金は貧困対策としても支持されていない。

Charap et al. (2013) はEIを指標としては用いていないが、66か国について1965-2010のパネルデータを用い、エネルギー消費の価格弾性値を $-0.3 \sim -0.5$ という結果を得たうえで、エネルギー補助金の廃止と貧困層にターゲットを絞ったセーフティーネット策の組み合わせで、大きな財政負担の軽減がなされるとしている。

林 (2017a) では都市化比率（中高所得国）のほかに、人口密度（低所得国）、コンフリクト要因＝ダミー変数（低所得国）とEIの改善との間に

有意な相関が認められた。エネルギーの消費効率が経済活動の集約度に影響を受けていることは、香港やマカオなど超高密度な経済がEIが最も低いランクに位置していることから、ある程度予想できる。いわゆる「コンパクト・シティー」政策はSDGs7の見地から積極的に評価できる可能性がある。

しかし、EIはある国が置かれた条件で大きく異なる。また、マクロレベルのEIは活動、構造、効率の複合的な要因で決定されるが、単純にエネルギー消費量をGDPで割っただけのEIからはその国のエネルギー効率改善の姿は見えてこない。国によって発展段階、それに伴う産業構造は多様であり、単純な横並びのクロスカントリー分析では、各国のエネルギー面での持続性を判定することはできない。経年分析においては経済発展に伴う産業構造の変化も勘案する必要がある。このため、活動、構造、効率の要素に分解して分析することが必要になる。

（2）EIの要素分解分析（Divisia Decomposition Analysis）

エネルギー消費は、単に消費効率だけではなく、どのような活動を行うかによっても左右される。エネルギーの消費を左右する要因は、経済活動の程度（Activity）、経済活動の構造（Structure）、気候・天候（Weather）、技術的効率性（Technical Efficiency）などである（IEA2020）（Dennehy 2014）。経済活動（A）が増加すれば、当然それは仕事量の増加からエネルギー消費の増大をもたらす。農業国が工業国に転換する、あるいは労働から資本への代替が進む過程ではエネルギーの消費が増大する。熱波や寒波が電力需要に影響を及ぼすことは日常経験からも理解できる。EIがエネルギー価格によって影響を受けることは言うまでもないが、構造（S）も大きな影響を与える。1973年に始まった第1次石油危機の後、日本では省エネルギー技術の導入が進んだだけでなく、エネルギー消費型の素材産業が海外に生産拠点を移すなどの構造転換が進んだ。

活動、構造、効率の関係は次式で示すことができる（EIA 2014, Heinen 2013など）。

$$E = \sum_i^n A \cdot \frac{A_i}{A} \cdot \frac{E_i}{A_i} = A \cdot \sum_i^n (S_i \cdot I_i)$$

エネルギー総消費 E

経済の総活動（付加価値、輸送人・キロなど） A

各分野の活動 A_i

分野ごとの活動の構造 S_i

分野ごとのインテンシティー I_i

この式にしたがって、エネルギー総消費の要素分解分析が可能になる。サブセクターごとに活動、構造、インテンシティーの項目を例示すると以下ようになる。暖房、給湯など人の生活に不可欠な部門では、活動は人口に比例するとみなしている。

表1 分野ごとの活動、構造、インテンシティー

部門	サブセクター	活動 (A)	構造 (S)	インテンシティー (I)
家計	暖房	人口	床面積／人口	暖房エネルギー消費／床面積
	給湯	人口	居住家屋数／人口	給湯エネルギー消費／居住家屋数
	調理	人口	居住家屋数／人口	調理エネルギー消費／居住家屋数
	家電製品	人口	所有者数／人口	家電製品使用エネルギー／所有者数
旅客輸送	自動車・バス 航空機・鉄道	乗客人・キロ	乗客人・キロシェア	使用エネルギー／乗客人・キロ
貨物輸送	トラック・鉄道・内航	貨物トン・キロ	貨物トン・キロシェア	使用エネルギー／貨物トン・キロ
製造業		付加価値	付加価値のシェア	使用エネルギー／付加価値
サービス業		付加価値	付加価値のシェア	使用エネルギー／付加価値
農業		付加価値	付加価値のシェア	使用エネルギー／付加価値

建設業		付加価値	付加価値のシェア	使用エネルギー／付加価値
-----	--	------	----------	--------------

（出所）IEA（2014）、Dennehy(2014)などに基づき筆者作成

この要素分解分析はIEAでも標準的なエネルギー効率の分析手法とされている。基準年と比較年の間の変化を見る方法、経年的な変化を追う方法などがある。

Wang（2013）は、要素分解分析を用い69か国の1980年と2010年の比較を通じて、技術進歩、資本蓄積、生産物の構造変化が経済全体のEIを改善する一方で、労働／エネルギー比率、空間的・時間的多様性がEIを悪化させるとする。Vogt, et al.（2014）は1995年から2007年の期間における40の主要経済における要素分解により、活動と構造の要素が経済全体のEIにどのように作用しているかの分析を行った。この結果、大部分の国においては技術進歩がEIの低下に貢献しているが、日本、米国、オーストラリア、台湾、メキシコ、ブラジルについては産業構造の変化の影響が強いとする。

要素分解分析のロジックに則ってSDGs7の目標設定と指標を再考してみたい。交通におけるモーダルシフトを例にとると、エネルギー消費の多い自家用車や航空機から旅客が鉄道に移転すれば、まず構造（S）面で旅客人・キロのシェアが変化し、経済全体EIを低下させる要因となる。また、その上で鉄道も含め運行単位当たりの燃費が改善すれば技術進歩を通じ交通分野のインテンシティー（I）の改善につながる。

一方、活動（A）はそれほど単純ではない。活動が増加することは豊かさにつながると考えられてきた。ビジネスにしても旅行にしても、移動は人々の生活を豊かにする要素である。観光は産業として確立しており、GDPを構成する重要な産業の一つである。コロナ禍で移動が制限されることは人々に大きな苦痛をもたらしている。ところが、移動しないことが豊かさにつながる一面があることも認識されつつある。遠隔会議の普及で、わずか1時間の会議のために一日かけて新幹線で東京・大阪間を移動することがいかに無駄であるかが社会の共通認識となりつつある。使用エネルギーと旅客人・キロが同率

で減少すれば全体のEIに変化はないが、もし、旅客人・キロの減少率の方が大きければ、輸送効率の改善 (I) がなかったとしても、EIは改善する。これは活動 (A) が生活の質の向上をもたらすかどうかについて問わなければならないということである。

要素分解分析を通じて、SDGs7に関しては、以下のような評価設問が浮かび上がる。

①活動 (A) は必要不可欠なものであり、アウトカムとしての厚生水準を向上させるものであるかどうか、②構造 (S) ないし構造の変化はその国の物的、人的資源の賦存状況から見て適切か、③マクロレベルのEIおよび分野別のIIは改善しているか、④全体としてエネルギーは消費、特に化石燃料などの枯渇性資源に由来するエネルギー消費は減少しているか (伸びが押さえられているか)。

この中で、最も重要と思われるのがアウトカムを生み出す活動の内容である。活動に関しては、製造業、サービス業などは付加価値を指標としている。また、輸送では旅客人・キロ、貨物トン・キロなどを指標としている。これらはアウトプットの指標であり、何が達成されたかというアウトカムの指標ではない。したがって、SDGs7の評価はエネルギーを使用した活動のアウトカムをどう定義するかという問題に突き当たる。

そこで、次節ではSDGs7が目標として掲げるクリーンエネルギーの普及が、どのようなアウトカムをもたらすかについて見てみたい。

3. 太陽光パネル・発電装置の導入による農村の変化

SDGs7の目標実現が人々の生活にどのような変化をもたらし、それらがSDGsの他の諸目標とどのような関係に立つかについて、具体的に見るため、半構造化インタビュー形式の調査をミャンマーとタンザニアで行った。この二つの地点は、1人当たりGDPが1,500ドル未満の国にあり、最大都市から300キロ前後の農村であることなど置かれた状況に共通点がある。両地点とも基本的に配電網が整備されていないオフグリッド地

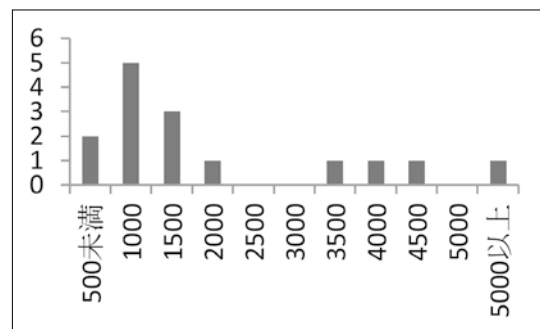
域であり、2010年代に急速に太陽光パネル・発電装置が普及した。導入時から時間が経過していないために、インタビューで回答が得やすいこと、筆者がこれまでも数次訪問し各種ヒアリングを行っていた地域で現地語に精通した研究協力者を得ることができたことなどが選択の理由である。調査項目はオフグリッド地域における家庭用太陽光発電装置の導入による家計と生活の変化、具体的には、導入内容 (設備容量)、導入費用 (月収の何か月分にあたるか)、用途、生活上の変化、住民の評価などであり、SDGs7の評価のためのプログラム・セオリーの構築を目指したものである。

(1) ミャンマー

ミャンマーでは2016年3月と2017年3月の2回にわたって、シャン州P村でインタビュー調査を行った。同村はオフグリッドであり、2010年以降急速に太陽光システムの導入が進み、2017年時点で普及率は95%～100%に達していた。2016年は予備的な調査で、対象数は15件、続いて2017年には本格的な半構造化インタビュー調査を行った。対象は同じく15件である。分析は2017年のデータを対象とした。

調査対象家計の年間所得分布は図3の通りである。平均値は1,878ドル、中央値は1,333ドルである。

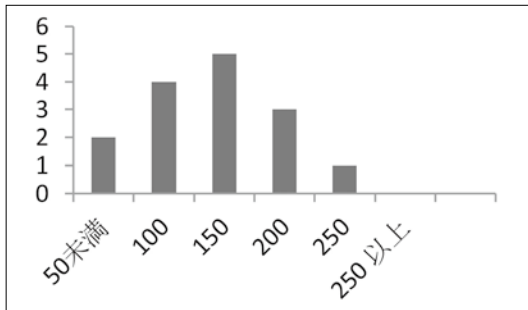
図3 家計収入 (ドル)



(出所) 筆者作成

導入された設備容量は図4の通りで、平均値131W、中央値130Wである。45W、100W、150Wなどが標準的なサイズである。200Wの設備容量を導入している家計は100W+150Wというような構成を並列でつないでいる。

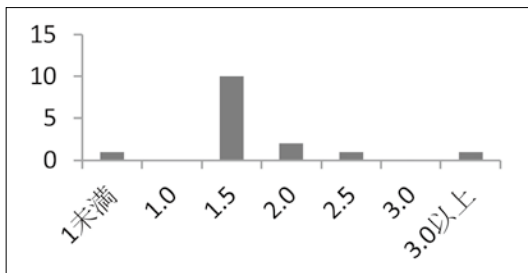
図4 導入設備容量 (W)



(出所) 筆者作成

Wあたりの導入コストは、平均値1.5ドル、中央値1.4ドルである（図5）。日本での太陽光システム価格は3.6ドル前後（2016-17年）であり、これに比べると相当安価であるが、これは中国製パネル、インバーター、タイ製バッテリーなどを使用した簡易なシステムであることによる。全家計が照明に使用している。TVは容量が小さい家計では使用していないケースもある。

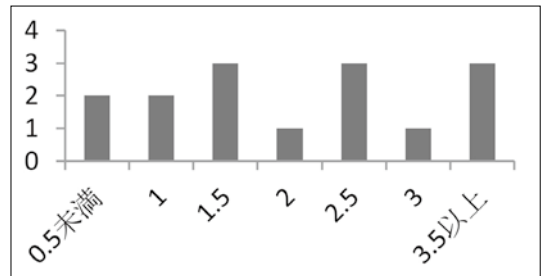
図5 Wあたりの導入コスト (ドル)



(出所) 筆者作成

導入費用が何か月分の月収に相当するかを求めたものが図6である。平均値は2.2か月、中央値は1.8か月であり、家計にとって大きな負担になったことがわかるが、ほぼ100%の普及率を見ると、負担以上の高い費用対効果を住民が認識していることが推定される。

図6 導入費用 (月収の月数)

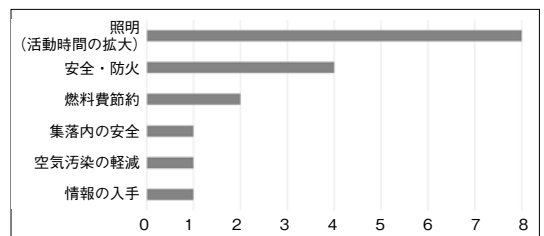


(出所) 筆者作成

インタビュー調査でも装置が高価なことに対する不満はそれほど多くは示されていない。太陽光装置の導入メリットとして回答者が指摘するのは以下の事項である（図7）。

- ①照明の使用（生活時間の延長、特に仕事と子どもの勉強の確保、街路の安全）
- ②室内大気汚染の防止
- ③火災の防止
- ④費用節約（燃料費）
- ⑤情報ギャップの縮小（TVからの情報、携帯電話の使用）

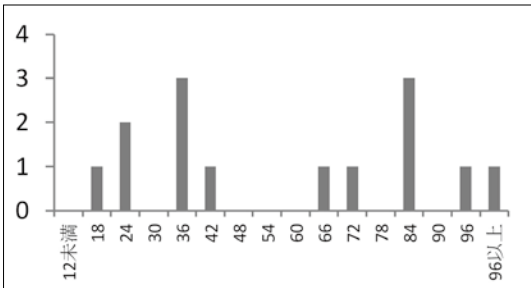
図7 導入のメリット (複数回答・回答数)



(出所) 筆者作成

燃料節約効果に関しては、1か月あたりのロウソク、ケロシン等の燃料費を導入費用で割った月数を図8に示す。平均値56か月、中央値45.9か月であり、燃料費だけに関しても約5年程度で、投資を回収できるとみていることがわかる。

図8 導入コスト／月間燃料節約額（＝月数）



(出所) 筆者作成

ヒアリングでは、太陽光システムに対する不満としては①曇天、雨天の日に使用できない（信頼性が低い）、②（中国製）パネルの耐久性の低さ、③設備を増強したいが資金がない、④グリッドにつなげてもらった方がよい（1例：高所得者）などの回答が見られた。全体としては、P村においては、太陽光システムは負担可能であり、住民の生活の質（QOL）の向上につながっている。

(2) タンザニア

タンザニアのデータ調査は、ミャンマーとほぼ同じ質問票を用いて、最大都市ダルエスサラームの西方250キロのM地区14村で行った⁸。回答数は全102⁹である。調査対象村の概要は表2の通りである。M地区も基本的にはオフグリッドで2010年代以降、急速に太陽光システムが普及している。

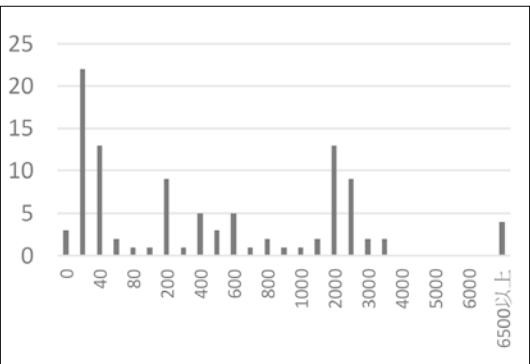
表2 調査対象村

	戸数	平均年収ドル
村		
A	1	158
B	2	1,267
C	2	7,392
D	1	1,584
E	1	1,584
F	1	2,640
G	1	1,584
H	1	365
I	1	317
J	2	1,267
K	21	395
L	11	2,424
M	16	2,147
N	41	45
平均値		973
中央値		167

(出所) 筆者作成

M地区は農業が所帯収入の大部分を占めている。家畜を保有している農民は比較的豊かであるが、農耕のみの農民は現金収入が限られている。老人のみの所帯も多く、（都市に住む）子ども・家族が生計を支えているケースが多い。家計収入の分布を図9に示すが、ばらつきがきわめて大きい。年間所得の平均値は973ドル、中央値は167ドルである。太陽光システムの普及率はほぼ50％である。収入が少ない家計では子ども、家族が費用を支出して太陽光装置を購入しているケースが多い。

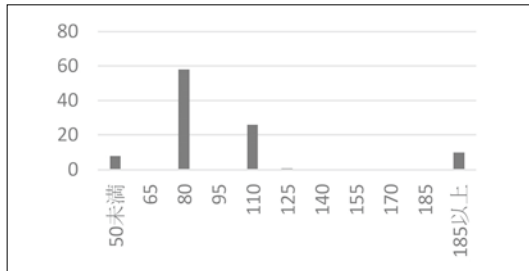
図9 家計収入（ドル）



(出所) 筆者作成

設備容量を図10に示す。平均値94W、中央値80Wでミャンマーより小規模である。

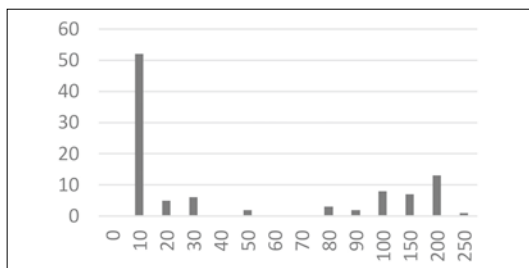
図10 導入設備容量 (W)



(出所) 筆者作成

導入コストの収入月数を図11に示すが、中央値は7.5か月、平均値は51.7か月である。図のようにM地区では30か月未満のグループと、負担可能なレベルとは言えない80か月以上のグループに2極化しているが、後者は上述のように都市に居住する親族に購入してもらっている。

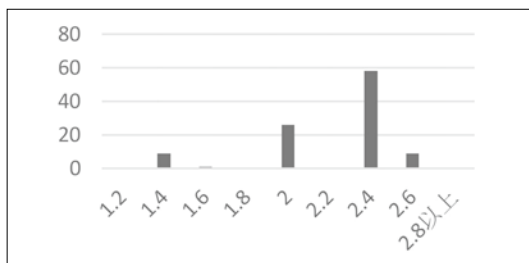
図11 導入コスト／月収 (月数)



(出所) 筆者作成

Wあたりの導入コストは平均2.1ドル、中央値2.3ドルであり(図12)、これもミャンマーよりも高額である。中国やインドなどの安価な製品を輸入するとしても、輸送コストの高さが大きな価格上昇要因であると推定される。

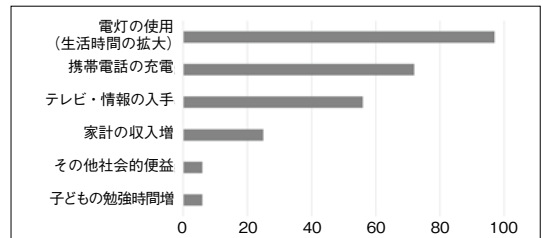
図12 Wあたりの導入コスト (ドル)



(出所) 筆者作成

村民が太陽光システム導入のメリットとしてあげる項目は図13の通りである。

図13 導入のメリット (複数回答・回答数)

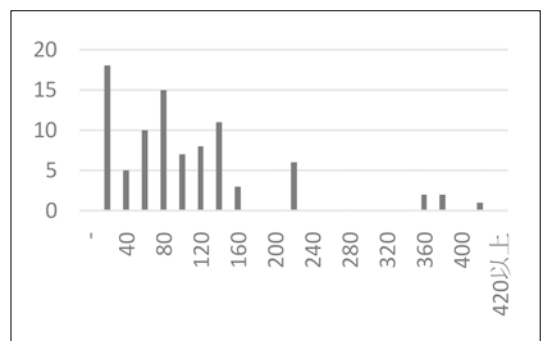


(出所) 筆者作成

これが示すように、電灯(照明)の使用(による生活時間の拡大)が最も多い回答となっている。25%の家計が収入増を回答しているが、これは作業時間の拡大と、普及率50%という状況の下で、携帯電話の充電をビジネス化している家計があることによる。

燃料節約効果に対する回答は図14の通りであるが平均95.8か月、中央値72.5か月であり、燃料費の節約だけでは費用の回収に8年ほどかかる計算になる。したがって、燃料費の節約よりは生活時間の拡大に大きなメリットを見出していることが推定される。

図14 導入コスト／月間燃料節約額 (=月数)



(出所) 筆者作成

(3) ミャンマーとタンザニアの比較

ミャンマーとタンザニアの比較を表3に示す。

表3 ミャンマーとタンザニアの比較

	ミャンマー	タンザニア
普及率	95～100%	50%
調査地の家計収入（ドル）中央値	中央値 1,333 平均値 1,878	中央値 167 平均値 973
導入コスト 月収月数	中央値 1.8	中央値 7.5
導入単価 ドル／W	中央値 1.4	中央値 2.3
KWH あたり発電単価（一日4時間、10年間使用）	0.095 ドル	0.158 ドル
燃料節約効果 回収月数	45.9	72.5
備考	(1) 満足度は高い	(1) 現金収入が乏しい家計が多い。 親類や家族の支援を受けてシステムを導入している (2) グリッドへの接続要望が大きい

(出所) 筆者作成

ミャンマーの場合、Kwhあたり発電単価（1日4時間、10年間使用とした場合）は9.5セント（約10円）にとどまっている。これに対し、タンザニアでは15.8セント（約17円）であり、先進国の太陽光FIT価格にかなり近い水準である（日本は2018年3月現在30円前後）。タンザニアの場合には一般の村民にとって、太陽光システムは負担可能と言うことはまだできない。これは所得水準と装置の価格に左右されるところが大きい。グリッドの拡大には多額の投資が必要であり、電源の問題（化石燃料による二酸化炭素の排出、大規模水力にともなう環境破壊、原子力発電のリスクなど）がある。太陽光システムの導入で、生活時間の拡大、情報ギャップの縮小など大きな生活の質の改善が見られることから、SDGs7が目標として再生可能エネルギーの普及を掲げているのは妥当であり、その中でも太陽光は大きなポテンシャルを持っている。ただし、低所得国でさらに普及させるためにはさらなる価格の低下が必要である。現在の数倍の出力が確保されれば、IH調理器の使用も可能になり、さらに有用性は高まる。ただし、近い将来、老朽化したパネルの処理の問題が生じる可能性がある。本調査ではパネル、バッテリー、インバーター

などの製造に要するエネルギーも含めた、総合的なエネルギー収支の検討まではできなかった。今後の課題である。

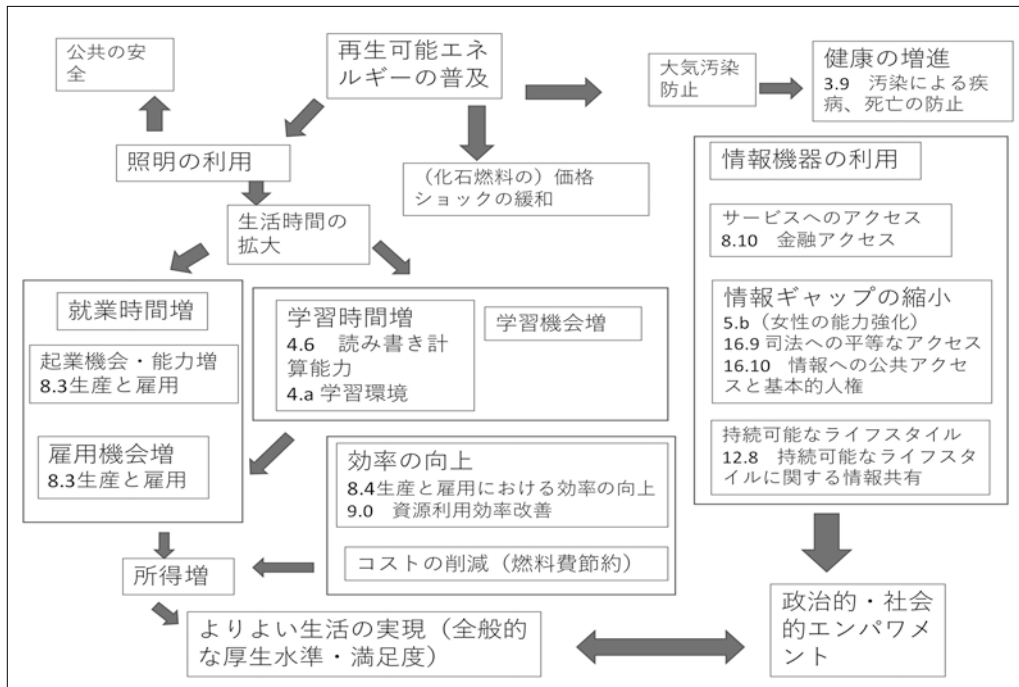
(3) ミャンマー、タンザニアの調査結果から想定されるセオリー

インタビュー調査などを通じて、SDGs7に関しては図15のようなセオリーを想定することができる。SDGsのゴール間で、因果関係、あるいはインプット→アウトプット→アウトカムの関係があり、目標横断的である。

まず、生活時間の拡大は夜間でも収穫物の加工や農具、漁具の整備を行うことを可能にし、就業時間増を通じて所得機会の拡大につながっている。これはインタビューでも確かめられている。この効果は、「生産と雇用（SDGs8）」に寄与している。

教育（SDGs4）との関係では、インタビューで子どもの勉強時間増を効果とする回答があり、学習機会増として把握することができる。電気が利用可能になり、テレビや携帯端末を通じて都市との情報ギャップが少なくなったことはインタビューを通じて、あるいはインタビュー外の一般的な会話でもよく聞かれた。これは、情報伝達の仕組みを構築することを通じて、ガバナンスの向上に寄与すると考えられる（SDGs16）。SDGsの他の目標を見てもエネルギーの供給を条件とするものは多い（例：SDGs12、生産と消費＝持続可能なライフスタイル）。その意味で、SDGs7はSDGs全体の中で極めて重要な位置を占めていることが改めて理解される。しかし、エネルギーの消費はインプットあるいはスループット（デリー1996）であって、それ自体が目的とするアウトプットでないことは明らかである。エネルギーの使用はそれ自体が目的ではない。人間が火を使用するようになったのは調理や採暖のためである。産業革命を経たのちは移動、運搬、計算などにエネルギーの使用が広がった。本質的な問題はエネルギーを使ってどのような仕事を行うかであり、同じ仕事をするのであればより効率的なエネルギーの使用が望ましいが、不要で無駄な仕事に費やすエネルギーをいかに効率的にしても意味はない。エネルギー効率を向

図15 再生可能エネルギーのセオリー



（出所）筆者作成

上させるとかえってエネルギー消費が増えるという「ジェボンズのパラドクス」があることを指摘しておきたい。EIを他の目標と切り離して評価を行うことは、このパラドクスを見過ごす危険性がある。

4. SDGs7のSDGs評価の中での位置づけ、およびSDGs評価の課題

SDGsの17の目標は一般には並列的に理解されているが、目標間には、複雑なインプット→アウトプット→アウトカムのロジカルなシーケンスがある。SDGs自体は目標相互間の関係について詳しくは触れていない。

SDGs7は、SDGsの目標（アウトカム）の達成を評価する中で、その達成のためにエネルギーの消費を増やすことになっていないか、同じ結果を出すために最もエネルギー消費の少ない手段が選択されているか、化石燃料などの枯渇性

資源の利用を抑え、再生可能エネルギーへの転換を進めているか、などの見地から評価すべきである。これはEIやエネルギー消費の弾性値のような単純な指標には還元できないため、かなり複雑になる。簡便な評価手法の開発は今後の研究課題である。

EIの指標としての問題点は視点を変えると分母として使用するGDPが経済のアウトカムを表す指標として適切ではないということでもある。この点はGDPを参照するSDGsの他の目標とも課題を共有する。GDPは評価学的に言えばアウトプットである。すべての活動が網羅されているのではなく、市場で貨幣によって取引されない財、役務、たとえば人々の自発的協力やボランティア活動はカウントされない。環境の悪化もマイナスカウントされない。枯渇性資源というストックを取り崩してフローに転換しても、フローの部分が国民所得統計に反映されるだけであり（デイリー 1996）、持続性を把握するためには不適切な指標である。

これまでもGDPに対する修正の試みは行われてきている¹⁰。資源ストックの増減などを織り込んだ欧州連合（EU）の「持続可能な発展指標（Sustainable Development Indicators）」、資源ストックの増減に生活の質を加えた経済協力開発機構（OECD）による「グリーン成長指標（Green Growth Indicators）」、防御的支出や余暇活動などの帰属計算を行うトービンとノードハウスによるMEW（Measure of Economic Welfare）（1972）やSMEW（Sustainable Measure of Economic Welfare）、さらに、これらを総合したISEW（Index of Sustainable Economic Welfare）などの提案が行われてきている。1993年、国連は国民経済計算体系（SNA）の改訂に際してサテライト勘定の一つとして環境・経済統合勘定（SEEA: System of Environmental-Economic Accounting）の導入を勧告した。これは環境関連の支出額（実際環境費用）や資産額（環境関連資産額）、さらに経済活動に伴う環境の悪化（自然資源の減少など）の「帰属環境費用」をSNAから分離して控除したものである。

これらの試みは、基本的にはGDPの延長線上にあるが、帰属計算の複雑さやあいまいさもあり、一般的に使われるには至っていない。

2008年のリーマンショック以降に検討が本格化したスティグリッツ委員会（経済成果と社会進歩の計測に関する委員会 The Commission on the Measurement of Economic Performance and Social Progress）の検討と提言は、全く新しい発想に拠るもので、経済活動のアウトカムを単一の指標で示すことは困難であるとし、複数の指標を用い、主観的幸福、ケイパビリティ（機会の選択）、公正な配分で生活の質を測定すべきであるとした（Stiglitz et al. 2018）。現在OECDが幸福度指数を公表している。まだ試行段階であるが、健康状態や社会とのつながり、ガバナンスなどの複数の生活の質および所得などの物質的な生活条件の指標を把握したうえで、母集団の平均値と下位集団間の差異を測り、さらに幸福のために資源として自然資本、経済資本、人的資本、社会関係資本のレベル（増減）を評価する内容になっている。普及が望まれるが、これもEIの分母とするような使い方は難しい。

GDPを代替するマクロ指標の確立は持続性の評価に不可欠であり、SDGsの成否を左右するものであると言っても過言ではないが、主観的要素も含まれることになると、クロスカントリー分析には必ずしもふさわしくないと思われる。使い方としては、一国のある時点からある時点までの変化を見ることによって、その国が持続性に近づいているかそうではないかを判断することが可能になると思われる。SDGs7の評価においては、これらのGDPを代替する指標の変化率とエネルギー総消費、化石燃料消費の変化率と比較することにより、生活の質（QOL）を低下させずにエネルギー面での持続性を高めているかどうかを判断することができるようになると考える。

5. 結語

SDGs7の評価の検討から見えてくるのは、まずSDGsの諸目標間の関係を整理しなければならないということである。SDGs7はそれ自体がアウトカムではなく多くの目標にインプットもしくはスループットとして関わっており、単独で評価するのではなく、他のアウトプット目標に関連付けて評価することが不可欠である。

SDGsは全体として「持続性」を十分にとらえていない。SDGsは依然としてGDPを基本とする従来からの思考にとらわれており、人間の限らない欲望を是認して「今の世代」も「将来の世代」も満足させることができるとの幻想に立脚している。SDGsの期間も1/3過ぎた今、次の目標に向けて物質的な「過剰」を見なおし、成長に頼らない社会システムを構築する段階に来ている。定常型社会（広井2001a；2001b）への移行は資源と環境の制約を考えれば必然である。人間の限らない欲望を見つめ「足ることを知る」人間観（安原2010）を目指すことも意義ある提案である。さらに、資本主義のシステムそのものに対する問いかけも不可避である（広井2015；斎藤2020）。その意味で、SDGsは本格的なパラダイム転換までの暫定的なものと考えらるべきである。

謝辞

本研究のうち、タンザニアのデータ収集は、科研費にJP16K13348の助成を受けたものです。データ収集にご協力いただいたGeorge Pindua 氏、貴重なコメントをいただきました査読者の皆様、アドバイスをいただいた皆様に深く感謝いたします。

注記

- 1 https://www.unic.or.jp/news_press/features_backgrounders/31591/ (2021年6月26日閲覧)
- 2 <https://dashboards.sdgindex.org/> (2021年3月1日閲覧)
- 3 World Commission on Environment and Development (1987)
- 4 BPによる。
<https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/oil.html>。
地域別に差があり、中南米は114、ヨーロッパが12である。(2021年3月11日閲覧)
- 5 資源エネルギー庁
<https://www.enecho.meti.go.jp/about/whitepaper/2018html/2-2-2.html>
(2021年3月11日閲覧)
- 6 2000年代の最初の10年間では「ピークオイル論が活発に議論された(レゲット2006)
- 7 エネルギー消費の増分をGDP増分で除した値をエネルギー消費の対GDP弾性値という。これが1を超える場合にはGDPの成長よりもエネルギー消費増の方が多く、マクロレベルでエネルギー消費効率が良好であることを意味する。
- 8 本件調査のデータ収集にあたっては現地在住の研究者George Pindua氏の協力を得た。
- 9 回答は103件得られたが、年収に他の回答と3桁の差があり、明らかに誤記と推定される1件を除いた。
- 10 以下は氏川 (2014) に拠るところが多い。

参考文献

朝野賢司 (2015) 「膨らむ再エネが貧困層直撃賦課金方式見直しをドイツ等で明らかになるFITの“不都合”」、Wedge Infinity (<https://wedge.ismedia.jp/articles/-/5282?page=2>) (2021年3月11日)

- 氏川恵次 (2014) 『環境・経済統合勘定の新展開』、青山社
- 斉藤幸平 (2020) 『人新生の「資本論」』、集英社
- デイリー H.E. (1996) 『持続可能な発展の経済学』、みすず書房
- デイリー H.E. (2014) 『定常経済は可能だ』、岩波書店
- 林薫 (2017b) 「開発の終焉 Development (開発・発展) からEnvelopment (整序) へ」、『東洋文化』、第97号、東京大学東洋文化研究所、169-196
- 林薫 (2017a) 「開発・発展の長期的持続可能性に関する考察—持続可能な開発目標 (SDGs) を意味あるものとするために」、『文教大学国際学部国際学研究叢書』、創成社、775-825
- 広井良典 (2001a) 『定常型社会』、岩波書店
- 広井良典 (2001b) 『グローバル定常型社会—地球社会の理論のために』、岩波書店
- 広井良典 (2013) 『人口減少社会という希望 コミュニティー経済の生成と地球倫理』、朝日新聞社
- 広井良典 (2015) 『ポスト資本主義 科学・人間・社会の未来』、岩波書店
- 村上由美子、高橋しのぶ (2019) 「GDP を超えて—幸福度を測る OECD の取り組み」、『サービソロジー』、6(4) : 8-15
- 安原和雄 (2010) 『足るを知る経済—仏教思想で創る二十一世紀と日本』、毎日新聞社
- レゲット, J (2006) 『ピークオイル・パニック 迫る石油危機 代替エネルギーの可能性』、作品社
- Charap, J., Ribeiro da Silva, A., Rodriguez, p. (2013). Energy Subsidies and Energy Consumption – A Cross-Country Analysis, *IMF Working paper* WP/13/112
- Dennehy, E. (2014). Moving from data to developing indicators, Instruction to the Energy Efficiency Indicators: Essentials for Policy making, Power point presentation at IEA Energy Training Week. (see IEA (2014))
- Filipovic, s., Verbić, M., Radovanovic, M. (2015), Determinants of energy intensity in the European Union: A panel data analysis, *Energy*, 92, 547-555.
- Heinen, S. (2013). Analyzing Energy Use with Decomposition Methods, Power point presentation at IEA Energy Training Week, <https://pdfs.semanticscholar.org/379e/5db6b0e7eb96e3bb84a54f0049078617cf5c.pdf> (March 31, 2021)

International Council for Science (2017). *A Guide to SDG Interactions: From Science to Implementation*.

<https://council.science/publications/a-guide-to-sdg-interactions-from-science-to-implementation/>

(March 31, 2021)

IEA (International Energy Agency) (2014). *Energy Efficiency Indicators Essentials for Policy Making*.

https://iea.blob.core.windows.net/assets/c41341f3-2149-4f59-a2e4-81c48bbc49be/IEA_EnergyEfficiencyIndicators_EssentialsforPolicyMaking.pdf

(March 31, 2021)

IEA (International Energy Agency) (2020). *Statistical Report, Energy Efficiency Indicators, 2020 Edition*.

<https://www.iea.org/reports/energy-efficiency-indicators-overview> (March 31, 2021)

Stiglitz, J., J. Fitoussi and M. Durand (2018). *Beyond GDP: Measuring What Counts for Economic and Social*

Performance, OECD Publishing, Paris

<https://doi.org/10.1787/9789264307292-en>.

(March 31, 2021).

Voigt, S., De Cian, E., Schymura, M., Verdolimi, E. (2014).

Energy intensity developments in 40 major economies: Structural change or technology improvement? *Energy Economics*, 41, 47-62.

Wang, C. (2013). Changing energy intensity of economies in the world and its decomposition. *Energy Economics*, 40, 637-644.

World Commission on Environment and Development (1987). *Our Common Future*. Oxford University Press.

(2021.7.13 受理)

付表 1

目標	指標
7.1 2030 年までに、安価かつ信頼できる現代的エネルギーサービスへの普遍的アクセスを確保する。	7.1.1 電気にアクセスできる人口比率。 7.1.2 汚染の少ない燃料と技術を利用可能な人口比率。
7.2 2030 年までに、世界のエネルギーミックスにおける再生可能エネルギーの割合を大幅に拡大させる。	7.2.1 最終エネルギー消費における再生エネルギーの比率。
7.3 2030 年までに、世界全体のエネルギー効率の改善率を倍増させる。	7.3.1 エネルギー・インテンシティ。
7.a 2030 年までに、再生可能エネルギー、エネルギー効率及び先進的かつ環境負荷の低い化石燃料技術などのクリーンエネルギーの研究及び技術へのアクセスを促進するための国際協力を強化し、エネルギー関連インフラとクリーンエネルギー技術への投資を促進する。	7.a.1 汚染の少ない燃料の研究とハイブリッドシステムを含めた再生可能エネルギーのための途上国への国際的な資金フロー。
7.b 2030 年までに、各々の支援プログラムに沿って開発途上国のすべての人々に現代的で持続可能なエネルギー、サービスを提供できるよう、インフラ拡大と技術向上を行う。	7.b.1 エネルギー消費効率向上のための投資の GDP 比率及び持続可能な開発のためのインフラ及び技術に向けた海外直接投資の GDP 比。

(出所) 外務省サイト (<https://www.mofa.go.jp/mofaj/gaiko/oda/sdgs/pdf/000101402.pdf> および <https://www.mofa.go.jp/mofaj/gaiko/oda/sdgs/statistics/goal7.html>) より筆者作成 (2021年3月21日)

付表2

ターゲット	相互作用	スコア	政策オプション
7.1 → 1.4	基礎的サービス、資源、資産へのアクセス改善。	+2	低廉なエネルギー、利用手段（改良型かまど）などの供給。
7.2 7.3 → 1.4	再生可能エネルギー導入に伴う貧困層が受ける価格ショック（貧困削減を阻害）。	0/-1	価格ショックに対する対応策。貧困層に対する支援。
7.2 7.3 → 1.4	再生可能エネルギーの普及による気候変動の緩和。気候変動によって被る貧困層の影響の緩和。	+2	総体的な二酸化炭素削減の取り組みが不可欠（効率性の改善や技術導入も含む）。
7.2,7.3 ← 6.1 6.2 6.4	安全な水の供給の要請にこたえるために海水の淡水化、下水の再利用などを行うと多量のエネルギーが必要になる。ただし、風力、太陽光などによる発電の変動を吸収することもできる。	+1/-1	淡水化や排水のリサイクルがエネルギー負荷を増大させないようにする。
7.2 → 2.1	農業における食料生産とバイオ燃料生産の競合。	0/-1	適切な土地利用計画、生産性向上。
7.2 → 2.3	食料生産+バイオ燃料+食料生産廃棄物活用による雇用の創出。	+2	女性、先住民、家族経営の農林水産業などに焦点を当てた多様化支援政策。
7.1 7.2 7.2 → 3.9	再生エネルギーの普及による大気汚染の防止。	+2	産業などにおける排出防止、室内大気汚染にさらされる人々への支援策。
7.3 → 3.4 3.6	アクティブ・トラベル（徒歩、自転車）による健康の増進。しかし、交通事故が増えるリスクがある。	+1/-1	アクティブ・トラベルがやりやすい環境を整える（歩行者・自転車道など）。
7.2, 7.3 → 6.1 6.4 6.5	再生エネルギーの普及による化石燃料の使用減少は、燃料生産過程での水使用（石油、ガス採掘に使われる水など）を減らすことができる。	+2	エネルギー政策や水政策の影響が、多国間の水配分に影響を及ぼさないよう配慮する。
7.1 → 8.3 8.5 8.6	エネルギーの普及によって、貧困層、特に女性や子供が仕事したり学んだりする時間を増やすことができる。	+1	エネルギーにアクセスできないことが、就業や学習にどの程度、どのように影響を及ぼしているかについては更なる調査研究が必要。
7.2,7.3 → 8.1, 8.4	二酸化炭素削減が経済成長にネガティブの影響を与える。	0/-1	慎重に考えられた、成長と環境のデカプリング政策を立案、実施。
7.2, 7.3 → 8.2 8.3 8.5 8.6 8.10	再生可能エネルギーやエネルギー節約技術の導入が新規雇用を生み出す。	+1	労働者の再訓練、中小企業、ベンチャー支援。金融へのアクセス。
7.2, 7.3 ⇔ 8.5	化石燃料関連産業での雇用喪失。	0/-1	
7.2 7.3 → SDG13	再生エネルギーの拡大は温暖化ガス削減の必要条件だが十分条件ではない。	+2	パリ協定の目標を達成するためにはすべての参加国の取り組みが必要。

(出所) International Council for Science (2017)より筆者作成

Evaluation of the “Goal 7” of the SDGs

Kaoru Hayashi

Bunkyo University
hayashi@bunkyo.ac.jp

Abstract

The 7th Goal of the SDGs is aiming at ensuring access to affordable, reliable, and modern energy for all. The difficulty to evaluate this goal is that use of energy is not an end objective of economic activities. Energy is being used to generate some outcomes, which must be targets of other goals though, the SDGs doesn't show any program theory of input, output, outcome relationship across the targets which requires inputs of energy. This article tries to construct a program theory by (1) analysing the concept of the Energy Intensity which 7th Goal deems as a key indicator for improving energy efficiency and by (2) analysing effects of introduction of solar power system, as recommended by the 7th Goal, in rural villages in Myanmar and Tanzania. The study suggests that while there are huge welfare gains for people through introduction of solar power system, the 7th Goal is best evaluated together with other goals and on how efficiently and sustainably these goals are achieved. This article further argues that basic indicators such as GDP should be improved to capture welfare and sustainability. It must be one of new challenges for goals after SDGs.

Keywords

Energy, Enhancement of welfare, Relationship between Goals, Energy Intensity, GDP

【研究ノート】

DAC 新評価基準の解釈と運用

江口 雅之

独立行政法人国際協力機構（JICA）

Eguchi.Masayuki@jica.go.jp

要 約

国連で2015年に採択されたSDGsの理念を反映すべくOECD-DACで約30年ぶりに評価基準が改定された。英語を原文とする新基準の定義で使われている用語は概念的なものが多いため、評価者は旧来の辞書的な固定概念にとらわれずに、用語の意味を正しく解釈した上で、適用するプログラムやプロジェクトの文脈に則し、適切に運用することが重要である。

本論では筆者が注目する新規に導入された用語の解釈について考察し、その運用の意義を解説する。

キーワード

DAC評価基準、公平性、人権、レジリエンス、援助協調

1. はじめに

「誰一人取り残さない」という理念の下に2015年にアジェンダ2030（SDGs）が国連で採択された。開発協力を取り巻くアクターの多様化や事業の複雑化を背景に、1991年以来約30年間に亘り改定されていなかったDAC評価基準の見直しは、SDGs採択を契機に本格的な改定の機運が高まった。SDGs採択後、国連評価グループは2016年にSDGs評価に関する会議を開催した。また、世銀評価グループ局長は現行の評価基準の課題を自身のブログでテーマ毎に発信し、評価関係者へ問題提起を続けた。2017年10月のDACハイレベル会合で、アジェンダ2030に沿ったプログラム評価基準の見直しが決議され、DAC下部組織の評価ネットワーク（EvalNet）が評価基準の改訂作業を担うこととなる。日本もメンバーのEvalNetでは、事務局が2018年にワークショップ

と公開コンサルテーションを実施し、現行基準への評価や改定基準案について広く意見を求めた。約700件に及んだコンサルテーションの意見では、約9割が現行の5つの評価項目（Criteria）の維持を支持しつつも、定義の明確化や見直し、基準新項目の追加提案が数多く寄せられた。

新評価基準の改定作業は、事務局及びメンバー国がDAC本部で一堂に会する定期会合の他、テーマ毎の数次にわたるオンラインによるウェビナーや電子メール、アンケートなどによる意見交換が繰り返され、2019年11月のEvalNet会合で最終案の合意に至り、同年12月のDAC会合で採択された。

その結果、従来のDAC評価項目の「Relevance（妥当性）」「Efficiency（効率性）」「Effectiveness（有効性）」「Impact（インパクト）」「Sustainability（持続性）」に、新たに「Coherence（整合性）」が追加されるとともに、各々の評価項目が再定義

された。改定の主要点は、SDGs理念の反映と援助協調の評価基準における明示的導入である。ただし、これらDAC基準の定義には概念的な用語も少なからずあるため、使用者により解釈と運用の違いが生じうる。また、解釈が不十分であれば、誤用し、不適切な評価を招くことにもなりかねない。

本論では、筆者のEvalNetなどの会議での経験も踏まえ、新基準として特徴的かつ解釈の差が生じやすい「Equity」「Human Rights」「Resilience」「Coherence」を取り上げ、各用語の概念を考察し、意義と運用上の留意点を解説する。これらが評価に携わる関係者にとって、新基準の適切な理解と運用改善の一助となれば幸いである。

図1 DAC新評価基準の概観図



(出所) OECD (2019)

2. 公平性とは？ EquityとEqualityは違う

新基準で明示された「公平性、ジェンダー平等、人権」は、SDGsの理念「誰一人取り残さない」を直截的に示したものである。先述の公開コンサルテーションのコメントでは、新基準として最多の提案となった概念は、Equity, Gender, Human Rightsであった¹。

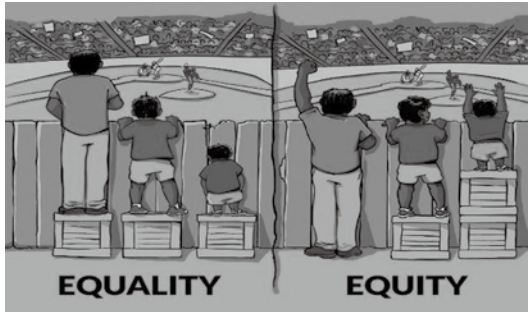
日本ではEquityとEqualityの訳語の「公平」と「平等」の違いを日常的に強く意識することは少ないが、両者には明確な違いがある。例えば、南米のチリでは約半世紀ぶりに憲法改正をすることが決まった。1970年代初めにクーデターで

社会主義政権を倒した軍事政権はネオリベリズムの経済政策を推進した。

軍事政権下で改定された憲法は「機会平等」の「自由」を推進する一方、社会福祉面からの「公平性」確保の視点が弱かったことが論点となっている。公平と平等は似て非なるものであり、時として対立する概念ともなる。

図2は公平性と平等の相違を示す一例であり、野球場の外から塀越しに試合を観戦しようとする身長差の異なる少年がいた場合に、同じ高さの踏み台を用意することはEqualityであり、異なる高さの踏み台を用意して視線の高さをそろえることがEquityである。Equityは機会の平等では

図2 EquityとEqualityの違い



(出所) Angus Maguire, Interaction Institute for Social Change

なく、均等な結果を保障するものである。

もし少年の中に車いす使用者がいれば、踏み台だけではなくスロープを設けることで公平性を確保することが出来る。その場合、スロープの整備には費用の増加を伴うが、費用便益分析の便益には反映されにくい。公平性の評価のためには、経済分析では測りえない固有の尺度が必要となる。

この公平性と平等の違いについて、筆者はブラハで開催された2019年の国際評価学会で英国の学者による両者の概念の説明を聞いて認識を深めたが、EvalNetの新基準の検討過程では取り立てて議論となったことはなく、当初より公平性という概念がこだわりなく用いられていた。欧州諸国のメンバーが多いEvalNetでは、公平性は当然のように定着していた概念だったのかもしれない。

新定義案の初稿案では、公平性は新基準が導入するAgenda 2030の主要素の1つとして認識され、「妥当性」における内容の適合性の範囲として、「有効性」における主要な評価視点として、「長期的変化」(のちに「インパクト」と変更)における不公平の是正の視点として、「Synergies (のちにCoherenceへ変更)」における公平性などを考慮した国際基準へのコミットメントとの政策的整合性の視点として登場した。

公平性を「有効性」の主たる評価視点とすることについては、例えば特定集団をターゲットとした技術移転プロジェクトのように、公平性への配慮が必ずしも求められない場合もある。そうした事例も考慮し、公平性は「有効性」の

主目的ではないが、目的達成と共に考慮すべきものとし、“including different result across group”が定義に追加され、プロジェクトの目的にかかわらず、社会の周縁層への影響について評価することを促す内容となった。

開発プロジェクトにおいては、公平性の視点の有無や程度により、評価の達成指標や評価判断が異なってくる。公平性の観点からは、例えば配電や上下水網の整備は、貧困層向けへのインフラ整備だけでは不十分であり、支払能力に配慮した料金設定までを検討する必要がある。インフラ事業では、徴収料金による投資及び運営維持管理費用の回収(Cost Recovery)が持続性の観点からは重要だが、公平性の視点を入れれば、富裕層と低所得層との、あるいは家庭向けと企業向けとの料金差別化や補助(Cross Subsidy)を進めることが必要になるかもしれない。

しかしながら、公平性の対象範囲については、ジェンダー、地域、社会階層など、様々であり、プロジェクトが準備できる予算で全てを網羅できるとは限らない。公平性に配慮をした追加的施策については、後年度の追加整備や同時期の代替的施策などを、プロジェクト形成段階において認識し、あらかじめ関係者間で方策を確認しておくことが重要となる。

ところで、公平性に配慮すべき対象には、社会的「弱者」が想定されるであろう。ただし、「弱者」は社会において相対的なものであり、障害者や貧困層がいつも「弱者」になるとは限らない。例えば、都市鉄道のエレベーターの設置に関しては、車椅子使用者への配慮が必要となるが、貧困世帯であっても健脚な者には必要としない。他方、配電事業の料金体系に関しては、貧困世帯への配慮を要するが、富裕層の車椅子使用者へ配慮する必要性は相対的に低い。公平性の配慮の対象と施策は、プロジェクト毎の裨益対象に応じて検討されることとなる。

3. Human Rightsの対象は？

人権の概念は幅広く、アブリオリなものから

国家による保障まで捉え方は様々であり、古くは英国のマグナカルタから、フランス革命のフランス人権宣言など、歴史的に進化してきたものである。第二次世界大戦のアウシュビッツやヒロシマなどの悲劇の反省を踏まえて策定された世界人権宣言は、戦争などの人々の争いを避けるために人類が創り出した最高の叡智とも評される。その理念に拘束性を持たせたものが1966年に採択された国連人権規約（「経済的、社会的及び文化的権利に関する国際規約（社会権規約）」及び「市民的及び政治的権利に関する国際規約（自由権規約）」）であり、今日の国際社会において最も広く共有され、尊重されるものである。日本政府も1977年に両規約を批准している。これら人権規約の中でもSDGsとの関連性が強いのは社会権規約であろう。別の言い方をすれば、人権はSDGs全般に通底している概念である。

現代社会における人権の考え方は国連規約で定められており、プロジェクトの計画時には配慮すべき対象を特定しておくことが実務的には必要となる。この場合、プロジェクトによる人権への影響は、プロジェクトの実施段階における非自発的強制移転のような場合には事態が表面化しやすいが、公共サービスのアクセスに関する地域や階層による差別などは可視化されにくく、意図せずに生じる場合もある。例えば公共サービスの支払い方法がクレジットカードに限定されると、銀行口座やクレジットカードを保有していない貧困層は、結果的にサービスから排除されてしまう。

また、人権概念の進化の歴史と同様に、その擁護されるべき対象と配慮の運用は定常ではないと考えるべきである。例えば、ITが進化する今日、インターネットへのアクセスに関しては、個人の権利として保障されるべきであると同時に、アクセス情報の第三者の監視と悪用による個人のプライバシー侵害の懸念も生じている²。これらは人権規約を定めた時代には到底想定しえなかったことであろう。また、2011年に国連人権理事会で承認された「ビジネスと人権に関する国連指導原則」では企業の人権尊重責任を定めており、プロジェクトに直接・間接に関わるサプライヤー企業の対応についても、大きな関

心に向けられている。開発協力を支える財・サービスの供給元において、児童労働や奴隷労働などの労働搾取や危険な環境下での労働による人権侵害が生じていないか、それら財・サービスを調達する国家や実施機関やプロジェクトを支える援助国が、どのように人権尊重の遵守を監督してゆくのか、今後益々その取組みが注目されるであろう。さらに、昨今の現代社会を大きく揺るがすコロナ禍の影響により、脆弱な状態におかれた人々の人権が侵害されるリスクが高まっている³。環境変化が早く複雑な現代社会においてこそ、人権への配慮は何かが必要なのかを、既成の認識に甘んじることなく、不断に考察することが求められる。

DAC基準改定の公開コンサルテーションでは、人権は公平性やジェンダーと並んで新項目案としての提案が最も多かった概念である。新定義案の初稿では、「公平性」と同様に新基準が導入するAgenda 2030の主要素の1つとして認識され、「妥当性」におけるプロジェクト・デザインの厳格性の視点として、「有効性」における包摂性で考慮すべき視点として、「長期的変化（のちにインパクトへ変更）」における不公平削減の視点として、「Synergies（のちにCoherenceへ変更）」における人権などの国際基準へのコミットメントとの政策的整合性の視点として、随所に登場した。また、改定議論の過程では人権の定義の明確化を求める意見もあった。

しかしながら、人権の視点を加えることへの賛意はあっても、人権を短期の評価指標として計測することは容易なことではない。そこで、新基準で「介入により実現した肯定的、否定的、意図的、非意図的な上位レベルの効果」と再定義され、「長期かつ広範囲で、福祉や人権、ジェンダー平等、公平性、環境など、より重要であり、システムや規範の包括的かつ永続的な変革的なもの」と解釈される「インパクト」の中で人権は上位目標として位置付けられた。

なお、DAC基準は、妥当性と有効性（妥当な計画に基づいた目標の達成度合い）や有効性と持続性（成果の継続の度合い）の関係に見られるように各項目の内容は相互に関連性を有している。人権侵害や人権配慮については、単にブ

プロジェクト実施の結果だけに着目するのではなく、計画デザインの段階、プロジェクトの実施段階、そして完了段階でプロジェクトが将来の人権侵害を誘発する可能性までを俯瞰することが重要である。

4. Resilienceは「頑丈」なだけではない

DACの新評価基準で新たに登場した用語の1つに、持続性に関するResilienceがある。近年、世界各地における自然災害やパンデミックへの対応にも関連し、介入効果の持続性を確保する重要な要素としてResilienceが認識されている。新基準では「持続性」の再定義に関連して、注釈部分でResilienceが挙げられている。

Sustainability: The extent to which the net benefits of the intervention continue, or are likely to continue. (OECD 2019) (介入の純便益が継続する又は継続する可能性の度合い。)

Note: Includes an examination of the financial, economic, social, environmental, and institutional capacities of the systems needed to sustain net benefits over time. Involves analyses of resilience, risks and potential trade-offs. Depending on the timing of the evaluation, this may involve analyzing the actual flow of net benefits or estimating the likelihood of net benefits continuing over the medium and long-term. (OECD 2019) (注釈：長期的に純便益を維持するために必要なシステムの財政的、経済的、社会的、環境的、制度的能力の調査を含む。レジリエンス、リスク、潜在的なトレードオフの分析を含む。評価のタイミングに応じ、純便益の実際のフロー分析や、中長期の純便益の継続見込みの計算を含め得る。)(下線は筆者による。)

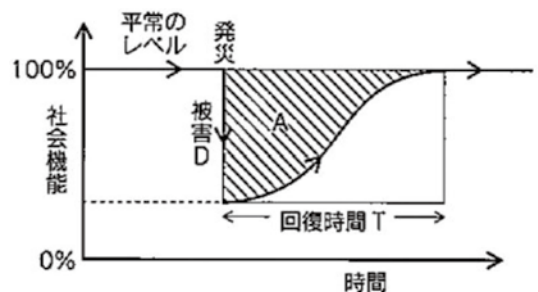
ResilienceはEvalNetで特に議論された概念ではないが、新定義が採択された後の日本語への定訳過程で、日本国内での解釈が使用者により不確かであることに筆者は気が付いた。

日本国内では、Resilienceは往々に「強靱性」と訳されることが多い。辞書の訳語は①弾性、

反発力、②変化からの回復力、③強靱>であるが、「強靱」は「頑健」というイメージが強いのではないだろうか。日本学術会議の災害に関する委員会では、「(レジリエンスの) もとからの意味は、「外部から力を加えられた物質が元の状態に戻る力」と「人が困難から立ち直る力」とされている。現在は「あらゆる物事が望ましくない状況から脱し、安定的な状態を取り戻す力」を表わす言葉として広く用いられている(日本学術会議 2014、p.24)」としている。

レジリエンスは、外的ショックへの耐性とダメージからの回復力の2つの概念を有するものであろう。図3は防災面でのResilienceを説明したものであるが、災害時の社会機能の低下を出来るだけ抑えたと共に、平常時の機能へ出来るだけ早く回復することがレジリエンスを計る尺度となっている。

図3 レジリエンスの意味



(出所) 河田恵昭 (2015)「災害復旧とレジリエンス再考」

世界経済フォーラムのGlobal Risk Report 2013では、Resilienceの定義について、工学分野では一般的な「ストレスがかかった後により速く跳ね返り、より大きなストレスに耐え、一定量のストレスによって妨げられることが少なくなる能力」という物理的解釈に留めず、社会システムの機能にも焦点を当て、「1) 状況の変化に適応する、2) 突然の衝撃に耐える、3) 以前の状態または新しい状態のいずれかで、操作の継続性を維持しながら、望ましい平衡状態に回復機能する能力」と定義した。更に、国毎のレジリエンスの構成要素として、5つのサブシステム

(経済、環境、統治、インフラ、社会)及び各システムの5つの要素(Robustness(頑健性)、Redundancy(余力)、Resourcefulness(問題解決のための資源の豊富さ)、Response(反応)とRecovery(回復))を考慮する必要があるとしている。

このように工学分野の従来のレジリエンスの解釈に加え、国際社会や国家システムへの危機対応の文脈で応用しているのである。他方、Resilienceについては、行政レベルでも部分的な理解による誤用も観られるようである⁴。今日、自然災害や感染症などのパンデミック危機に対して、ダメージを皆無にすることは困難である。これらの危機に対して予知と予防により頑健さを備えるだけではなく、被害のダメージを可能な限り縮減し、ダメージから早期に機能を平常時に戻すあるいはそれ以上に回復させる(Build back better)ための対策を想定し、用意しておくことがレジリエンスの本質と言えよう。

個々の開発プロジェクトを取り巻くリスクには、自然災害によるもの、人的なもの(意図的・非意図的)、国内外の政治・経済環境の変化によるものなど、様々であるが、危機管理の視点からリスクを把握し、備えておくことが、プロジェクトの効果持続を担保するためには不可欠である。

DAC新基準が2019年12月のDAC本会合で採択されるのと時をほぼ同じくして、世界的な新型コロナウイルスのパンデミックが始まり、EvalNetでは新型コロナウイルスのパンデミック対応に関する評価タスクが立ち上げられ、短期・中期のセクター別の分析や横断的対応、過去のパンデミック対応の評価と教訓が収集・分析され、提言作成に向けた作業が開始された。現在も将来も絶えることのないパンデミックに対して、社会的損失をいかに縮減できるか、そして社会機能をいかに早期に回復できるかを、真の意味でのレジリエンスの視点として評価することが重要である。

5. 縦の糸と横の糸、Coherence

Coherenceは新基準改定の目玉である。Coherenceが新基準として追加された背景には、上述のSDGsの理念とは別途、近年の開発協力を取り巻く環境変化と、プロジェクトの担い手の多様化と複雑化に対し、介入の適切な評価を行うためには、従来のDAC評価基準では不十分という関係者間での問題意識があった⁵。

Coherenceの辞書の訳語は、< ①結合力、密着、②(論理などの)統一、一貫性、③【物理】可干渉性 > である。DAC新基準ではCoherenceはInternal CoherenceとExternal Coherenceに分けられた。Coherenceは組織内外の統合性、整合性、一貫性という要素を含んでいる。政策・プログラム・プロジェクトの一貫性が縦の糸であるとすれば、介入する側の組織内外の連携が横の糸であろう。これら縦の糸と横の糸によって丈夫な布が織られれば、線から面へ援助効果が確かなものとなる。

しかしながら開発援助の現場では、プロジェクトで多様なアクターが介入する場合に、適切に調整されていれば相乗効果や補完効果を生むが、各アクターの無関心や利害関係により、投資資源の無駄やプロジェクト実施上の障害・混乱が生じうる。そして、その影響を最も直接的に被るのは被援助国の市民である。

国際協力の世界では、新評価基準の改定議論が本格化する以前から、人道支援などのドナー間の協調介入の調整(整合性、一貫性、重複排除、相乗効果など)が高い関心事項であった。援助協調や援助の質・効果に関する議論は、調和化ローマ宣言(2003年)、パリ宣言(2005年)、ア克拉行動計画(2008年)、釜山宣言(2011年)で議論・合意されてきたものである。各援助ドナーからの調整を経ない支援は、被援助国側に十分な管理・吸収能力がなければ混乱を生み、ドナー側の貴重な資源も無駄となりかねない。そうしたことは理屈の上ではわかっている、援助協調は時として多大な調整労力を要するため、強いインセンティブが働かなければ意識的に取り組まれないことも多い。

これまでCoherenceや援助協調の視点は、既に

独自の基準として設定している国・機関もあったが⁶、今般、DACの基準として正式に位置づけられたことは、援助協調の精神が、評価の世界においても国際標準として認知・運用されるという意味において画期的である。世界各地で分断や対立の増加が危惧される現代の国際社会においてこそ、積極的な「協調」や「連携」の意義がある。

Coherenceは複合的な意味を持っており、日本語では一語で表すことが容易ではないが、例え「整合性」と訳されても、単に各ドナーの援助方針に不整合が見られないとか、ドナー間における重複が無いから事足りているとするのではなく、被援助国の便益が最大化することを目指して、複数アクターによる結合的な協力による相乗効果を求める姿勢が必要である。

ところで、DACの評価基準はDACメンバー国の検討議論によって定められ、運用において考慮・尊重されるものである。一方、中国などの新興ドナーや開発途上国はEvalNet内での議論には参加しておらず、DAC基準に縛られることもない。その観点から、援助協調を包括的かつ効果的なものにするためには、非メンバー国へのアウトリーチの取組みが不可欠である。その点において、今般の改定議論で中国の主導で発足したアジアインフラ投資銀行（AIIB）がオブザーバーとしてEvalNet会合へ招待されていたことは重要な試みであったと思われる。そして今後、新基準で追加された援助協調の精神を確かなものとするためには、多数のプロジェクト実施と評価の実績を有する国連や国際金融機関、JICAらによる新興ドナーや開発途上国への積極的な評価知見の共有と関与が一層必要である。

6. おわりに

今回のDAC評価基準の改定では、開発プロジェクトの評価においてSDGsの理念を積極的に取り込むことが意図された。本論で列举した用語の解釈はその一端に過ぎないが、言葉の旧来の辞書的な訳語にとらわれず、適用目的や対象に合わせて解釈・運用することが重要である。

DAC評価基準の定義で用いられる用語は国際的な共通理解のための「共通言語」としての性格を有しているが、それらはドナーの拠出額や債権放棄額のように解釈が確かなものばかりではない。各国の言語文化を背景として概念的な理解の幅・相違が生じることは、ある程度は止むを得ないが⁷、各国評価者が用語の真意を的確に理解し、適切に運用することが重要である。

SDGsの理念を反映すべくDACが改定した新基準による評価は、評価の意義である「評価は現実世界の活動であり、重要なことは、政策やプログラム、実践に変化をもたらし、早かれ遅かれ人間の生活の状況を改善するような広がりをもつことである。（ロッシ2005、p.340）」⁸ということの具現化を希求するものであろう。

新型コロナウィルスが猛威を振るい、人々の生活スタイルを大きく変化させたパラダイムシフトに直面している今日の我々にとって、将来の政策やアクションのよすがになるのは、客観的な評価や検証により導き出される教訓なのではないだろうか。

（本稿は筆者の個人的見解であり、組織の意見を代表するものではない。）

注記

- 1 新基準提案の上位よりequity, gender, human rights 57件、partnerships, synergies 49件、inclusiveness, stakeholder engagement 49件、co-ordination 26件、environment, climate change 25件、design rigour 22件（Summary of Consultation Response, EvalNet, November 2018）
- 2 Berti(2018)は「基本的人権の定義が常に変化しているため、人権法は常に変化している。例えば私たちの生活のデジタル化が進めばネット上のアクセス権やプライバシー権が必要かもしれない」と指摘。
- 3 名越（2020）は「サプライチェーン上の非正規雇用者（パート、移民労働者など）は、正規雇用者比較して、労働搾取や労働機会喪失のリスクが高い可能性がある。また、このような状況下で、失業者が新たな職を求める結果として、人身取引や、強制労働

- のリスクにさらされることなどが想定される」と指摘。
- 4 中央防災会議防災対策実行委員の河田関西大学教授は、Disaster Resilienceは縮災であり、減災と人間社会の能力（共助力や早期回復時間）からなるものであり、National（Community）Resilienceを「国土強靱化」としたのは誤訳と指摘している。
- 5 Hider（2018）
- 6 国際開発センター（2010、p.44）は、IADB、CIDA、DFIDがCoherenceを導入済みであり、DAC 評価5 項目の他にパリ宣言への対応として「マネジメントの原則・パリ宣言の順守」（Management Principles/Adherence to the Paris Declaration）、「調和・調整」（Harmonization and Coordination）などの項目を含めている機関がある（CIDA、DFID）」と指摘。
- 7 国際開発センター（2010、p.46）は、「未だ用語の使い方がばらばらで、それぞれの機関が言う有効性がどのレベルの評価を指しているのか、相互に理解するのが難しい状況にある」と指摘。

参考文献

- 河田恵昭（2015）「災害復旧とレジリエンス再考」、関西大学社会安全研究センター
- 小稲義男編「新英和大辞典 第五版」（1980）、研究社
- 国際開発センター（2010）「平成22年度「政策レベルのODA評価（手法・体制）」に係る調査」報告書、外務省
- 国連広報センター「世界人権宣言70周年：人権の考え方」、
https://www.unic.or.jp/activities/humanrights/document/bill_of_rights/universal_declaration/
- 名越正貴（2020）「新興国におけるサプライチェーン上の人権課題と日本企業への影響」-JICA緒方研究所セミナー（EY新日本有限責任監査法人講師の配布資料）
- 西谷修（2019）「ロジェ・カイヨワ 戦争論」、NHK出版
- 東日本大震災復興支援委員会 災害に対するレジリエンスの構築分科会（2014）「提言 災害に対するレジリエンスの向上に向けて」、日本学術会議
- ロッシP. H.、リプセイM.W.、フリーマンH.E.（2005）「プログラム評価の理論と方法」、日本評論社
- Berti, Benedetta（2018）. *70 year universal declaration of*

- Human Rights - #standup4humanrights*, UNIC（国連広報センター）.
- DAC EvalNet（2018）. *Summary of Consultation Response*
- Hider, Caroline（2018）. *Rethinking Evaluation*,
<http://ieg.worldbankgroup.org/sites/default/files/Data/RethinkingEvaluation.pdf>. The World Bank.
- OECD（2002）. *Glossary of Key Terms in Evaluation and Results Based Management, first edition*.
- OECD（2019）. *Evaluation Criteria: Adapted Definitions and Principles for Use*.
- OECD（2020）. *Thoughtful Evaluation: A Guide to Applying Criteria, draft 0.4*.
- World Economic Forum（2013）. *Global Risk Report 2013*.

（2021.7.6 受理）

Interpretation and application of DAC New Evaluation Criteria

Masayuki Eguchi

Japan International Cooperation Agency

Eguchi.Masayuki@jica.go.jp

Abstract

The OECD-DAC has revised new evaluation criteria for the first time in about 30 years to reflect the SDGs philosophy adopted by the United Nations in 2015.

Since some of the terms newly introduced in the new evaluation criteria in English are conceptual, the evaluator should apply them after correctly interpreting the significance of these terms without being bound by the old dictionary-like stereotypes.

In this paper, taking up four terms (Equity, Human Rights, Resilience, Coherence) that the interpretation of terms is likely to be of ambiguity, insufficiency, or misunderstanding, the author shows the points to be noted together with the actual use of each term.

Keywords

DAC Evaluation Criteria, Equity, Human Rights, Resilience, Aid coordination

【実践・調査報告】

新規 SDG 事業構想時の評価活用： アフリカの次世代企業家育成事業の事例報告

長尾 眞文

一般財団法人 国際開発センター

Nagao169@gmail.com

要 約

新規事業を立ち上げる際の事前評価は、通常の場合策定された事業計画の内容や実行可能性を評価する目的で実施される。しかし、SDGs関連事業の中には分野横断的取組や判定の難しい質的成果を目指すプロジェクト等、新規の立ち上げに複雑な要因が絡むケースが増えており、計画の段階から評価的思考を明示的に応用する事前評価実施の必要性が次第に顕著になっている。本稿は、筆者自身が関与したアフリカの次世代企業家育成事業の新規立ち上げに際し、J・オウエンの事業前評価（Proactive evaluation）のフレームワークを活用して計画と評価が同時進行する事前評価を実施した事例の報告である。事例から得られた教訓は、事業企画の初期の段階からステークホルダー対話を重視することの重要性と聞き取りやフォーカス・グループといった質的方法とワークショップのような集合的データ創出手段を効果的に組み合わせる可能性である。

キーワード

評価事例報告、事前評価、事業前評価、アフリカの人材育成、企業家人材

1. はじめに

評価の世界では、ある程度の規模の社会的事業を新規に立ち上げる際に事前評価を行うことは常識と化している。標準化された公立学校建設、インフラ、職業訓練のような事業でフィージビリティ・スタディを実施して新規計画の実行可能性チェックをすることはもちろん、貧困削減や社会的弱者救済のようなソフトの事業申請案にもプロジェクト・デザイン・マトリックスやログ・フレームの形で評価の観点を反映することが必須要件となっている。しかし、2015年9月の国連総会決議「持続可能な開発のための

2030アジェンダ」によるSDGs時代の到来は、次の三つの要因で評価の事前的な活用の必要性和重要性をさらに一段と高めた。

第一の要因は、2030アジェンダ下で立ち上げられた無数の新規SDGs事業間の競争的資金獲得の圧力である。申請者が資金提供者に対して説得力のある事業案件を提示するには、案件形成過程での評価的思考の明示的な導入とその申請案への反映が不可欠で、それが事前評価の活用を誘引したと考えられる。

第二の要因は、SDGs事業の取組の複雑化である。持続可能性を追求するSDGs事業には、水供給と公衆衛生、農業振興と栄養改善といった複

合目的や分野横断的アプローチに特徴づけられる案件が多い。教育・保健・ジェンダーの格差解消のように、取組の有効性や成果の質的把握の難しい事業も少なくない。しかも複雑性のゆえに、類似の先行案件も限られる。そこで事前に事業の成否の可能性について丁寧な検証を行う必要が生じる。

第三の要因は、「誰一人取り残さない」のサブタイトルに象徴される2030アジェンダのステークホルダー重視の方向性である。実施事業は必然的に参加型となるが、その円滑な運用には事業計画作成の段階で各ステークホルダーの意見や希望を把握し、事業の達成目的や具体的取組の設定に反映する必要がある。

新規事業の立ち上げのプロセスは、通常事業の計画策定と計画内容の評価の2段階で考えられ、前者は計画論、後者は評価論の課題と区別して考えられがちである（JICA 2004、p.129）。しかし、SDGs事業に関わる前掲の三要因の示唆するところは、計画策定段階の複雑化のためこの区別が不明瞭で計画と評価が一体化しつつあることである。例えば、SDGs関連の競争的資金提供者は、単に資金を拠出するのではなく、自らのSDGs指針や支援方法に関するガイドラインを持っている。従って新規事業計画者は企画策定の段階から資金提供者による評価を意識しなければならない。前例のない分野横断的な新企画や質的成果を目指す事業案については、事業計画者自身が評価方法を案出する必要を迫られる可能性がある。多くのSDGs関連事業では主要なステークホルダーが事業の実施に参加するだけでなく、成果の評価にも参加することが想定されており、その双方が事業策定段階で考慮される必要がある。要するに、SDGs関連の新規事業の策定では計画者が評価者を兼ね、評価的思考を十分に活かすことが求められるのである。そこでは、事前評価は出来上がった新規事業計画の実行可能性を検証するのではなく、むしろ新規事業企画形成プロセス全体の有効性を担保する役割を持つ。

本稿では、以上のようなSDGs関連事業の文脈で、筆者が実際に係わったアフリカの次世代企業家育成のための新規事業企画について事前評

価の事例報告を行うこととする。以下、次の第2節で事前評価事例の経緯と評価課題の説明、第3節で新規事業企画の具現化に向けた評価活用の概念的フレームワークの提示、第4節で事前評価事例の実施プロセスの説明、そして第5節で事例からの教訓のまとめを行う。

2. 事前評価事例：経緯と評価課題

本稿で事前評価の事例報告をするのは、国連大学サステイナビリティ高等研究所（以下、UNU-IASと略称）が2009年～2018年に実施した「アフリカの持続可能な開発のためのリーダー養成プログラム」¹（以下、英文名称のイニシャルを取ってESDAと略称）の一環として、2016年から新たに立ち上げを検討した若手企業家育成企画である。新企画の母体に当たるESDAはアフリカ5か国・8大学の連携協力事業で、本稿の筆者はUNU-IASの客員教授としてそのコーディネーターを務めていた。ESDAは2013年以来「総合農村開発」、「都市化」、「鉱物資源開発」の3分野で大学院修士課程レベルの専門人材育成プログラムを共同実施しており、表1で示すように2015年時点で在学生総数は全体で75名を超え、卒業初年度に当たった同年、8名がプログラムを修了した。修士課程の設置に手間取っている大学もあったが、ESDA全体としては次のような諸特徴を備えたアフリカの大学間共同修士プログラムの定着を意識するようになった。

- 1/ アフリカ有力大学によるプログラムの共同オーナーシップ
- 2/ 各大学の長所・強みの組み合わせと優秀教員・教材の共用による改善プログラムの実現
- 3/ 開発現場の重視、問題解決アプローチの多用による実践的プログラム内容
- 4/ 国連機関、日本・北欧の大学による実施支援
- 5/ ICT利用の遠隔教育手法の活用
- 6/ アフリカ以外の学生への門戸開放

表1 ESDA修士プログラムの在学学生数と卒業生数：2015年

(単位：人)

修士プログラム	大学	在学学生： 2015	卒業生： 2015
アフリカの都市化	ナイロビ大学（ケニア）	8	0
	ケニヤッタ大学（ケニア）	0	0
アフリカの鉱物・鉱業資源開発	ケープタウン大学（南アフリカ）	21	1
	ザンビア大学（ザンビア）	11	0
アフリカの持続可能な農村開発	ガーナ大学（ガーナ）	0	0
	クワメ・エンクルマ科学技術大学（ガーナ）	24	8
	開発研究大学（ガーナ）	5	3
	イバダン大学（ナイジェリア）	7	0
合計		76	12

(出所) ESDAプログラム・ファイル

2015年9月の国連総会によるアジェンダ2030の採択を契機として開かれたESDA参加大学のプログラムのレビューでは、「次の課題領域」としてアフリカの持続可能な工業化の推進を担う次世代の企業家育成が提起された。これはSDG目標9「産業と技術革新の基盤をつくろう」を意識したものであったが、その背景には、2000年以降アフリカ諸国の経済成長が加速傾向に転じたこと²、アフリカ各国政府自身が意図的な工業化を優先政策課題に掲げ始めたこと（African Development Bank, et al. 2017）、そしてそれを受けてアフリカ連合が2015年にアフリカ各国政府の超長期戦略に関する総意として採択した「アジェンダ2063」で持続可能な工業化の推進を中核に据えるという大状況の展開があった（African Union Commission 2015）。ESDA参加大学の間でアフリカの工業化の推進を担う次世代の企業家をいかに育成するかについて理解や経験を共有している訳ではなかったが、この課題をESDAの四つ目の修士プログラムとして結実させるという暗黙の了解が成立した。

ESDA事務局を担当していたUNU-IASでは、2016年からアフリカの工業化と企業家の育成の課題について本格的な文献調査を開始した。そこで分かったことは、開発経済論者の間でもアフリカの工業化を現実的な政策課題として捉える研究が盛んになっていること（Page 2013；Szirmai et al. 2013）、アフリカの豊富な若年労働者の存在

が「人口ボーナス」の形で製造業の成長を支える可能性が重視されていること（Newman et al. 2016）、東アジアの工業化経験がアフリカに持つ意味合いへの関心が強いこと（Khan 2011；Ohno & Ohno 2011）、国内市場組織の整備や経済体制の運用に関する学習の必要性（Noman et al. 2011）やアフリカの地域経済統合を含む国際競争力強化の課題（African Development Bank 2015）が政策レベルでの関心と呼んでいることであった。また、経済の発展と企業家の役割に関するレビューからは、長期的な工業化プロセスの展開には企業、産業、経済活動でリーダーシップを発揮する企業家群の存在が不可欠であるにもかかわらず、工業経験の浅いアフリカでは未だ企業家層が十分に形成されておらず、工業化の政策的推進には継続的に企業家人材を輩出するメカニズムの創設が必須要件となること（Naude 2011；Acs and Naude 2012）が判明した。

ESDAでは以上の文献調査による検討に基づき、アフリカの次世代企業家人材育成の新規プログラム化を図ることを決定したが、若手企業家人材育成のニーズがどの程度あるのか、ESDA参加大学も含めてアフリカの大学がこの分野で実際にどのような関連プログラムを提供しているのか等々、企画案作成に不可欠な基礎的データが欠落していた。そのため企画の具体的内容は未確定のままであった。幸いにもアフリカ開発銀行や南アフリカ政府の研究支援基金（National

Research Foundation)等のドナー組織が若手企業家人材育成のテーマに強い政策的関心を持っていたことから、新企画の具現化に協同で当たることになった。協同の条件には、企画案の提示に加えて十分な裏付けデータの準備や企画案の評価の枠組みの提示、さらにこのテーマ自体の重要性に関して周知を図るための出版やネットワーク活動の実施等が含まれた。結果として、事前評価の課題として浮上したのは、計画作成と企画評価の双方を合体する形で、ドナー組織、大学、そして最も重要な学生に対して説得力のある新企画案を策定することであった。

3. 新規事業企画の具現化に向けた事前評価活用の概念的フレームワーク

以上の極めて包括的な新規事業企画の具現化の課題に対して、本事例では事前評価の方法としてOwen (2006) の‘Proactive evaluation’を活用することにした。原著の用語の‘proactive’は「前向きの」、「先取りする」、「事前対策となる」の意である。新規の事業企画に関する意思決定に判断材料を提供するという事前評価の一般的な役割の中で、特に企画の初期段階で系統的な評価作業を起動させ企画形成プロセスを前進させる評価との意味が込められている。但し、本稿ではこの方法を事業が存在する前の評価と解釈して「事業前評価」と訳すことにした³。

事業前評価は、事業担当チームが事業形成前のグリーン・フィールドの状況で構想の事業化を図り、具体的な企画案を作成するのに評価を活用するものである。事業前評価の典型的な活

用例は、学校教育現場で既存のカリキュラムが時代や環境条件の変化により陳腐化したとの判断からカリキュラムの改訂を考える時である。そこで必要とされる作業は、陳腐化の意味する欠陥を埋めるニーズの確定、関連カリキュラム研究の最前線の情報・知識の取得、他校のモデル・カリキュラム実践経験からの学習等の判断の鍵となる情報の収集・分析で、最終的にそのような評価調査の結果を総合して望ましい新カリキュラム案をまとめることである。

表2は、事業前評価を活用する際の概念的フレームワークを標準的な事前評価手法の一つであるフィージビリティ・スタディと対比する形で提示したものである。事業前評価もフィージビリティ・スタディも根源的な目的は同じで、新規の事業企画に関する意思決定に判断材料を提供することである。しかし、評価対象は極めて異なる。事業前評価の場合に対象は前例案件の少ない新事業領域の構想段階の案件であるが、フィージビリティ・スタディの場合は先行事例のデータや経験に基づいて作成された事業計画である。典型的評価設問で比較すると、事業前評価が事業のニーズや課題の把握について問い、関連する先行的取組やベスト・プラクティスの有無を探り、関係する調査研究の示唆を求めるのに対して、フィージビリティ・スタディでは市場、技術、財務、環境、事業実施等の要件に関する定型的な設問に対するデータやシミュレーションの答えが要求される。事業前評価の関心がいまだ構想段階にある事業企画を取り巻く環境的条件の理解にあるのに対して、フィージビリティ・スタディは新規事業計画の実行可能性に焦点化する。

表2 事業前評価の概要

	事業前評価	比較：フィージビリティ・スタディ
目的	新規の事業企画に関する意思決定に判断材料を提供すること	新規の事業企画に関する意思決定に判断材料を提供すること
評価対象	比較的前例案件の少ない新事業領域の構想段階の案件（例えば、分野横断的な農村開発、新分野の人材育成）	比較的豊富な先行事例のデータや経験に基づき作成された新規事業計画（例えば、道路、港湾施設等のハード・インフラ、製造業プラント）

	事業前評価	比較：フィージビリティ・スタディ
典型的評価設問	・この事業企画にはニーズがあるか？ ・課題の把握は十分か？ ・ステークホルダーの関心は？ ・この分野のベスト・プラクティスは？ ・類似の選取的取組は？ ・関連調査研究の示唆は？	・事業計画が的確に記述されているか？ ・事業の提供するサービスに対する市場条件が将来予測も含めて適切に把握されているか？ ・事業の技術的側面の検討は代案も含めて適正に成されているか？ ・事業は財務的に実行可能か？ ・環境への影響は十分考慮されているか？ ・事業の実施体制、作業スケジュール等に問題はないか？
評価対象事業の現況	・事業開始前	・事業開始前
主要なアプローチ	・ニーズ・アセスメント ・先行研究調査結果の統合 ・ベスト・プラクティスの検証	・先行類似案件の検証 ・技術的、経済的シミュレーション・スタディ ・環境アセスメント
データ収集方法	文献サーチ、アンケート、現場訪問・聞き取り調査、フォーカス・グループ、デルフィ法、ワークショップ	文献サーチ、二次統計データ、現地調査、聞き取り調査、フォーカス・グループ、デルフィ法、ワークショップ
成果物	・新規事業企画案	新規事業計画
実施上の特徴	① 内部的活用 ② 探索的实施 ③ 多角的な方法による多様なデータ収集 ④ フォローアップの重要性	① 外部の実施 ② 分野ごとに確立された手法の活用 ③ 課題の要請に対応する定型的数据収集 ④ 新規事業の実行可能性に関する独立的判断

(出所) 事業前評価についてはOwen (2006) / フィージビリティ・スタディについては筆者

事業前評価は新しい評価の試みであり、定型的な評価設計方法があるわけではない。評価者が実際に取り得るアプローチとしては、事業前評価ではニーズ・アセスメント、ベスト・プラクティス（ベンチ・マーク）、先行調査研究の結果の統合が挙げられ、総じて探索的である。それに対して、フィージビリティ・スタディは多くの場合既に確立された多様な手法があり、その組み合わせで設計が決まるが、調査は検証的である。

実際に特定の事業構想あるいは企画案について事業前評価を設計するための最初のステップは、この構想・企画で取り上げる問題・課題を特定し、その背景にある環境条件についての的確に記述することである。前述したように、事業前評価の選択は前例の少ない革新的な事業の可能性の探求が前提となっている。このステップでは、新たな構想・企画をめぐる固有な状況の

正確な把握が不可欠である。次いで第二のステップは、構想・企画案の事業化に向けて問われるべき評価設問の選定である。表2の典型的評価設問では先行事例やベスト・プラクティスに触れているが、当然のことながら検討対象事案が革新的であればある程先行経験・調査は限られ、不確実な要件の割合が拡大する。また強い革新性は現実の状況の否定あるいはそれからの脱却を意味し、ある種の価値判断を内包している。従って評価設問の選定に当たっては、可能な限り多角的な視点から設問を用意することを心がけると同時に、事案に関係するステークホルダーの意見を丁寧に聴取することにより設問の妥当性を確保することが重要である。第三のステップは評価アプローチの決定である。評価設問と同様に革新性の高い事案については特定のアプローチに固執せず、多様なアプローチを併用することが適当であろう。最後のステップはデ

ータ収集方法を含む評価調査の具体的取組の設定である。多角的な視点からの事案の検討、多様なアプローチの並行的採用は評価調査自体を複雑にする。評価調査の組立については最初から確定してしまうのではなく、調査の進行に合わせて頻繁に振り返りを行い、形式的に評価調査を改めつつ柔軟に対応することが適当であろう。

- (1) 企画の着想
- (2) 企画案作成方針の決定
- (3) 企画案裏付け評価の実行案作成
- (4) データ収集・分析
- (5) ステークホルダー対話
- (6) 評価のまとめ
- (7) 評価結果の活用

4. 事前評価事例の実施プロセス

本稿で報告する事例は、2016年から2018年にかけてUNU-IASがアフリカの企業家人材育成事業の企画案作成のために実施した事前評価である。その実施の具体的プロセスは下記のステップを踏んで行われた（表3参照）。

以下では、各ステップで実施された具体的作業の内容と結果について順を追って簡単に説明する。

(1) 企画の着想

ESDAでは隔年ベースで参加8大学が集まって共同実施プログラムの検討を行っていた。2015年10月の会議は2030アジェンダの関連での振り返りが主要テーマであった。議論がアフリカの

表3 アフリカの次世代企業家育成事業: 事業前評価の展開

評価のステップ	具体的作業	作業結果
企画の着想	・SDG を契機とするアフリカ専門人材育成事業（ESDA）の振り返り（2015年10月） ・アフリカの若手企業家人材育成の課題に関する文献調査の開始（2016年1月） ・アフリカの若手企業家育成の新機軸に関する探索的ワークショップの開催（ケープタウン、2016年3月）	・予備調査の実施に関する事業内合意形成 ・アフリカの若手企業家数名との意見交換を目的とするワークショップで研修訓練関心を確認
企画案作成方針の決定	・方針作成のための内部検討作業（2016年4～6月） ・複数のドナーと新規 SDG 関連事業開発調査への資金提供について交渉 ・ESDA 執行委員会による意思決定（クマシ、2016年7月）	・アジアの工業化経験の共有を軸とするアフリカの若手企業家人材育成の方向性 ・ドナー、ビジネス大学院を含む複数のステークホルダーとの協働による事業化方針 ・複数のドナーによるデータ収集活動支援の確保
企画案裏付け評価の実行案作成	・オウエンの事業前評価コンセプトの採用 ・多様な関連データの多角的収集と複数のステークホルダーとの収集データの共有・共同アセスメントによる漸進的・発展的評価の実行可能性の確認	・事業前評価実行計画の作成 ・ESDA 執行委員会の実行案への同意
データ収集 1	・アフリカの若手企業家との対話ワークショップの開催（ナイロビ、2016年8月）	・アフリカの若手企業家 30 名、ビジネス研究者 25 名との対話から企業家育成活動に関する具体的提案

データ収集 2	・ アフリカの主要ビジネス大学院の訪問・聞き取り調査（2016年11月～2017年4月）	・ アフリカの9か国・14大学のビジネス大学院の訪問・聞き取りによる情報収集及び大学間協力による共同企業家育成事業の実施の可能性に関する意見の聴取
データ収集 3	・ アフリカの主要ビジネス大学院の企業家教育カリキュラムの比較研究（2017年）	・ アフリカの3か国・9大学のビジネス大学院の企業家教育カリキュラムの検討に基づく現状と課題の把握
ステークホルダー対話	・ 主要ビジネス大学院研究科長ワークショップの開催（プレトリア、2018年4月）	・ アフリカの10か国・14大学のビジネス大学院の研究科長が、若手企業家数名、ドナー機関代表数名を交えてアフリカの若手企業家育成の課題について討議、ビジネス大学院ネットワークを形成し、課題別ワーキング・グループによる共同調査の実施について合意
評価のまとめ	・ 評価報告書の作成 ・ 並行的に実施した研究の成果物出版 ・ ESDA 執行委員会による検討（東京、2018年10月）	・ アフリカの若手企業家人材育成に関する新事業企画書の作成 ・ 研究書（オープン・ソース出版） <i>Youth Entrepreneurship and Africa's Sustainable Industrialization</i> (Spears Media Press, 2018) ・ ESDA 執行委員会の承認
評価結果の活用	・ 新規事業企画案を内容とする助成申請書を作成 ・ ドナー機関に助成申請を提出し、資金的支援の交渉開始（2019年）	・ ドナー機関の内部的事情で交渉が長引くうちに、新型コロナの蔓延でドナー機関の資金提供に関する優先順位が変更され、企画案は保留状態 ・ 企画案の代替的実行の可能性の模索の開始

出所： 事業前評価についてはOwen（2006）/ フィージビリティ・スタディについては筆者

若年労働者の雇用の課題に集中する中で、雇用創出効果の大きい工業化に注目が集まり、ESDAプログラムで実施中の農村開発、都市化、鉱物資源開発の3修士プログラムとのテーマ的な補完性、大学が学生を対象として実験的に運営する企業スタートアップ・インキュベーターの経験、ビジネススクールの現役ビジネスマン対象の企業家養成プログラム等について意見が交わされた。事前準備を伴わない議論であったが、SDG目標9「産業と技術革新の基盤をつくろう」の追い風効果もあって、ESDA参加大学は若手企業家人材育成プログラムの可能性を検討することで合意した。それを受けて、プログラム事務局を務めるUNU-IASは即座に課題理解のための文献レビューを中心とするデスク・スタディを開始

した。同時にESDA参加大学のビジネススクールのネットワークを活用して、実際に企業活動を行っているアフリカの若手企業家から起業の経験や研修ニーズに関する聞き取りも行った。

企画の着想を強く後押ししたのは、南アフリカ政府の研究助成機関のNational Research Foundation (NRF) であった。NRFは、アフリカの若手企業家育成のテーマで本を出版することを条件に、このテーマで小規模の探索的ワークショップを開催することに資金を提供してくれた。2016年4月に南アフリカ・ケープタウン市で、ESDA参加大学の一つでアフリカ有数のビジネススクールを擁するケープタウン大学と国連大学が共催し、アフリカの若手企業家数名とビジネス研究者を招いて問題理解の深化のための対話

ワークショップを開いた⁴。そこで得られた成果は、南アフリカとケニアから参加した若手のスタートアップ企業家が抱えている経営の課題について詳細な事例検討を行えたこと、そしてビジネススクールによる企業家研修や技術的・経営的支援の可能性について確かな展望を獲得できたことであった。また参加したビジネス研究者とアフリカの持続可能な工業化に向けた若手企業家育成の課題に関する本を出版するプロジェクトを立ち上げることもできた。

(2) 企画案作成方針の決定

ESDA事務局によるデスク・スタディと探索的ワークショップのポジティブな結果を受けて若手企業家育成に関する新規事業の企画案を作成することになり、その方向性も含めて企画案作成の基本方針を定めることが次のステップであった。このステップは基本的に事務局の内部作業で進められ、最終的にESDA執行委員会により決定された方針には、次の三点の特徴があった。

第一の特徴は、アフリカ各国に存在する三つのタイプの企業家人材育成機会— (i) 内外NGO、援助機関等によるプロジェクト・ベースの研修・訓練、(ii) 大学による学部生対象の企業・起業に関する講義・演習、(iii) 大学院による専門的経営・起業・企業家教育—のうち、ESDAの先行プログラムと統合的な大学院レベルの若手企業家育成プログラムの創出を目指すとしたことである。但し、アフリカの大学院レベルのビジネススクールは大学内で強い独立性を維持する傾向があることから⁵、新規企画はESDAプログラムの単純な外延的拡大として捉えることは難しく、新たな大学間ネットワークの形成を必要とすることが分かっていた。

第二の特徴は、工業化や企業家育成がアフリカにとって比較的新しい課題であり、新企画は大学が新分野創出の役割の一端を担うことを意味することから、若手企業家、企業、政府、ドナー機関等マルチ・ステークホルダーとの協働による事業化の方向性を追求することが不可欠としたことである。実際にこのステップでは、ドナーによる継続的支援の確保にかなりの努力を注入し、前述の南アフリカのNRFやアフリカ開

発銀行(AfDB)との連携が確定した。

第三の特徴は、アフリカの持続可能な工業化の推進に向ける新企画の軸にアジアの工業化経験の共有を据えることにしたことである。これはアフリカの工業化に関する文献レビューから新たな視点として浮上した方向性であるが、それを決定的にしたのはAfDBがアフリカの持続可能な発展を支援するためのアフリカーアジアの知識の共有・共創を主要テーマとすることを条件に銀行内の日本信託基金からの助成を決定したことである。

(3) 企画案裏付け評価の実行案作成

この段階で、新企画作成にJ・オウエンの事業前評価コンセプトを採用することにした。前例の少ない新たな分野で説得力のある新企画案を作成するには、作成プロセスに意図的・明示的に評価の思考を応用する必要があった。それは、具体的には若手企業家人材育成に関する多様なデータの多角的収集と複数のステークホルダーとの収集データの共有・共同アセスメントを軸として漸進的に課題の理解を深め、有効なプログラム介入案を組み立てる作業であった。複数のドナーとの継続的対話は極めて有効であった。例えば、AfDBの助成契約には結果ベースのログ・フレームが含まれていて、いくつか定量的な数値目標を明示していた。しかし、企画自体の発展的志向や人材育成プログラムの質の漸進的な向上の可能性を強調し、ドナー・スタッフ自身によるデータ収集活動やステークホルダー対話への参加を通じた検証で実施プロセスと達成成果の確認をするよう説得した。

(4) データ収集・分析

① アフリカの若手企業家との対話ワークショップの開催

事業前評価のデータ収集の取組は、予備的調査段階で形成したビジネス研究者ネットワークを活用した若手企業家人材育成事業に関する共同研究の実施であった。目的は二つで、第一は若手企業家が起業段階で直面する経営課題、特にスキル・ギャップの問題を明らかにすることにより人材育成事業のニーズを把握すること。

第二は既存の若手企業家育成・研修の先行的取組について調査することであった。その実現の方法は若手企業家と企業研究者の合同課題探求ワークショップを対話型で開くことであった。ワークショップは、2016年8月にナイロビでアフリカの若手企業家約30名、若手研究者25名の参加を得て開催された。ワークショップには若干名の日本人参加者も含まれていた。

このワークショップで得られた主なデータは、アフリカの若手企業家自身が感じている経営面、技術面の課題とそれに関連する支援ニーズに関する事例報告であった。アフリカの企業の実態調査をしている若手研究者からは報告に関するコメントや実施中の企業支援プログラムのデータが得られた。また日本人参加者との議論や交流、さらにはワークショップ・プログラムに組み込まれた日本企業や企業家に関する講義を通して、アフリカの若手企業家や研究者の日本やアジアに対する関心が明らかに増幅したことが観察された。

ワークショップ終了時には、参加した若手起業家・研究者の総意として、(i) 若手企業家の活動とイノベーションに関する情報を集約するための共同ネットワークの創設、(ii) 若手企業家と研究者が協働するための企業インキュベーション・センターの拡大と増設、及び (iii) アフリカの有力ビジネススクール間の連携協力による若手企業家人材育成プログラムの設立の三点から成る提案がなされた⁶。

この対話ワークショップの結果は、同時期にナイロビで開かれていた第6回アフリカ開発会議(TICAD VI)のメイン・セッションの一つで報告された。このセッションは、ESDAの提案に基づき、国連大学、ナイロビ大学、ケニア政府、AfDBの共催で「アフリカ若手企業家フォーラム」として組織された。このフォーラムは、AfDBの総裁やケニアの青少年担当大臣の講演を聴くことがメインの目的であったが、同時に若手企業家と研究者の提案をアフリカ各国の政府関係者に提示してフィードバックを取得するという副次的な目的を備えていた。ESDA事務局の実質的な関心は、このフィードバックを新企画に対する政治的支援に関する感触のデータとして捕捉

することにあった。実際に対話ワークショップの結果の報告に対する反応は好評で、企画案作成の推進圧力になった。またアフリカの若手参加者がTICAD VIに参加することで、ワークショップで芽生えた日本に対する関心を後押しする効果もあった。

② アフリカの主要ビジネス大学院の訪問・聞き取り調査

ESDA事務局は、アフリカの9か国・14大学のビジネス大学院を訪問して、教育プログラム内容に関する聞き取り調査を行った。ほとんどすべての訪問先で聞き取りの相手はビジネス大学院の研究科長であった。事前に大まかな訪問目的を知らせるほかは、アンケート的な質問リストを用意したりせず、各訪問で約2時間を使って対話的にデータ収集した。会話の中で意識的に発した質問は次の4問であった— (i) ビジネススクールの社会的役割をどう考えるか、(ii) 大学内でのビジネススクールの役割についてどう考えるか、(iii) 若手企業家の養成に関するニーズについてどう考えるか、そして (iv) 他大学のビジネススクールとの協力の可能性はあるか。主要な結果を要約的にまとめると次の通りであった。

- 1/ ビジネス大学院の主流は少数の大規模企業社員を対象とする社会人修士プログラム
- 2/ 教育内容は主として経営・管理理論と経営計画作成
- 3/ ほとんどのビジネス大学院が農村、インフォーマル・セクターを度外視
- 4/ SDGs課題に対する関心は稀薄
- 5/ 学内で他研究科から孤立
- 6/ ビジネス大学院間で優秀な学生と協力企業の取り込みのための競争
- 7/ アフリカの工業化を担う企業家人材育成プログラムの可能性に対する関心

もちろん訪問した14校の中には、社会的関心を前面に出して、大学院プログラムの中に中等教育未修了者のためのビジネス研修プログラムを創設したり、都市の貧困層が居住する地域で

起業活動を支援したりするビジネススクールもあった。また、僅かであるが、SDG目標9を明示的に掲げて積極的にイノベーション教育を推進しているビジネス大学院もあった。最後に、ビジネス大学院はお互いをライバル視する姿勢が目立つが、若手企業家人材育成のような新規の構想であれば多くが協働ベースで参加する余地はありとした。

③ アフリカの主要ビジネス大学院の企業家教育カリキュラムの比較研究

事業前評価のもう一つのデータ収集作業は、アフリカの主要ビジネス大学院の教育カリキュラムの比較を行うことであった。その主たる目的は、アフリカのビジネス教育のベスト・プラクティスの見当をつけることで、特にそこに新規事業を起こす企業家育成の要素がどのように組み込まれているかを見極めることであった。既に訪問して聞き取りを行った14校の中から地域的バランス、公立・私立の別、相互評価をもとに6校を選び、MBAプログラムに的を絞ってウェブサイトで得られる情報による比較検討作業を行った⁷。その結果得られた理解は次の3点であった⁸。

- 1/ アフリカで高評価を得ているビジネス大学院プログラム（特にMBA）は経営管理、金融・会計、ソフト・スキルのいずれかを強調する内容となっている。
- 2/ カリキュラム形成の主要な関心はアフリカ経済の現状でビジネスが直面している短期的課題にあり、アフリカの持続的開発に関わる問題意識は比較的希薄である。
- 3/ 企業家人材育成に対する関心は、個別企業組織の設立・存続のための企業家訓練・育成が主で、長期的に産業・経済を牽引するための企業家リーダー創成を志向するものではない。

④ データ収集作業のまとめ

以上のデータ収集作業で明らかになったことは、アフリカの次世代企業家育成企画について、若手企業家の側では極めて強い関心と経営的・

技術的支援に対するかなりの期待が観察されたのに対して、アフリカのビジネススクールの反応は必ずしもポジティブではなかったことである。調査を実施するESDA事務局がビジネススクールの現状理解を優先し、企画や構想の可能性について前がかりに話すことを意図的に控えたので、これはある程度事前に予想できたことであった。アフリカの大学の多くが経営的、財政的に危機的状況にあるにも拘わらず、ビジネススクールは総じて安定的な運営を継続しているだけに、現状維持の姿勢が目立つようであった。それに対するESDA事務局の対応は、既に訪問したビジネス大学院のネットワーク化を図るとともに、次に触れるステークホルダー対話を通して改革志向の明らかな少数のビジネス大学院が牽引役を務めるように働きかけることであった。

(5) ステークホルダー対話

以上の一連のデータ収集作業の結果を踏まえ、UNU-IASは2018年4月にNRFとAfDBの支援を得て南アフリカのケープタウン大学大学院ビジネススクールおよびプレトリア大学GIBSビジネススクールと共催でアフリカ10か国の15ビジネス大学院の研究科長を招いて対話ワークショップを開催した。主催3大学とドナー2機関の事前協議で、この会議の目的は、「ビジネス大学院の連携協力によりアフリカの持続可能な工業化を担う企業家人材育成事業を実施することの必要性と実行可能性について検討し、可能であればその実現のための協働作業の枠組みを作成する」こととした。

2日間にわたる会議には、参加ビジネススクールの多くが慎重な姿勢で臨んだ。その理由として挙げられたのは、企業家育成ニーズは多種多様で各国のビジネススクールの対応も一様ではないこと、アフリカ全域をカバーするプログラムの検討には必要とされる企業家育成状況や現行の学習指導方法に関する情報が不足していること、そして個々のビジネススクールの置かれている状況が極めて異なること等であった。結果として、まず参加ビジネススクール間で共同研究を実施し、ビジネススクールの現状と取り組むべき課題に関するより深い理解を共有する

必要があるとの認識で一致した。その上で各ビジネススクールが実際に企業家人材育成事業の実施で共同歩調をとるための環境整備に向けて、教員が大学の枠を超えて参加する課題ごとの「実践コミュニティ（Community of Practice / CoP）」を組織すること、アフリカの企業や企業家についてケース・スタディを実施すること、アフリカにおける企業家教育のあり方に関する共同研究と教育実践アプローチの共有を行うことで合意した。

(6) 評価のまとめ

2016年7月のESDA執行委員会の決定を皮切りに実施したアフリカの若手企業家人材育成事業企画案の事業前評価は、ほぼ2年後の2018年8月までに一連のデータ収集活動とステークホルダーとの協議を終え、評価のまとめ作業に入った。その内容は探索的・漸進的であった評価調査プロセスそのままに、実行に移す形で評価作業の完成を見た部分的な企画案と引き続き継続評価作業を要する課題表記の部分を併記するものとなった。まとめ作業で重視したのは、評価調査の前半で実施したアフリカの若手企業家との対話と後半のビジネス大学院研究科長とのステークホルダー対話であった。また調査の全行程に参加した少数の日本側の若手企業家と研究者の視点、さらにはドナー・スタッフによる継続的モニタリングとそれに基づく協議も最終的な企画案の検討に有用であった。最終的にまとめた企画案の主たる項目は下記の通りであった。

① ビジネス大学院教員の「実践コミュニティ」 設立の企画案

ビジネス大学院研究科長ミーティングの提案を受けて、全参加校の若手教員が下記の3課題について個別の実践コミュニティを形成すること。

- 1/ 企業家教育用のケース・スタディ集作成のための研究開発
- 2/ 教員の指導力改善プログラム
- 3/ 企業家支援目的の中核インキュベーション・ネットワークの形成

2018年8月と11月にプレトリア大学で全実践コミュニティが参加して、AfDBに対して助成申請する目的で実行企画案の作成を行った。

② アジアーアフリカ企業家交流プログラム

日本／アジアとアフリカの若手企業家がお互いの知識と経験を交換し、共同事業開発の可能性を模索すること。複数のドナーにプログラム実施の支援要請をするために申請案づくりを行った。

③ 企業家教育用教材・資源の共同活用

短期的に企業家教育用教材の共有により教育単価を引き下げる企画、中長期的に年1回ベースで企業家教育カリキュラム改善研究フォーラムを開催する企画の継続的検討を行うこと。

④ アグリビジネスのインキュベーション・ネットワーク形成

ビジネス大学院研究科長会議に参加した数校がアグリビジネス分野のインキュベーション・プロジェクトの推進に強い関心を持っていることから、さらに新たな企画案追加の可能性を探ること。

2018年10月に東京の国連大学で開かれたESDA執行委員会以上を内容とする事業前評価の結果の報告がパワーポイントを用いてなされた。また事前評価と並行してビジネス研究者ネットワークをベースとして実施された共同研究が完成し、全12章から成る研究書として2018年10月にオープン・ソースで出版された（Kapfudzaruwa, et al. 2018）。

(7) 評価結果の活用

ESDA執行委員会による事前評価の報告の承認を受けて、ESDA事務局は2019年に入って前掲の企画案、特にビジネス大学院教員の「実践コミュニティ」設立の企画案とアジアーアフリカ企業家交流プログラム申請案の2案を完成させ、複数のドナー機関に助成申請を行った。しかし、2018年末を境にESDAプログラム事務局を南アフリカ・プレトリア大学に移転したこともあって、ドナー機関との支援交渉が遅れがちになった。

さらに2020年に入ると新型コロナの蔓延によりドナーの資金提供に関する優先順位が変更され、ESDAの申請案は複数のドナーで保留状態になった。2021年になってようやくアジア・アフリカ企業家交流プログラムの実施に対して、東芝国際交流財団とミシガン州立大学からの助成が得られることとなり、本稿で事例報告した事前評価の結果が活かされることになった。

5. 事例からの教訓：事前評価の有用性と課題

SDGs関連事業の特徴の一つは分野横断的取組や判定の難しい質的成果を掲げるプロジェクト等多種多様で、その評価にも多様なアプローチの活用が必要となる。本事例報告では、SDG目標9「産業と技術革新の基盤をつくろう」の後押しもあってアフリカ諸国の主要な政策関心として浮上しつつある持続可能な工業化の課題に着目し、その実現を担うべき若手企業家人材の育成企画案の事前評価について報告した。先行事例が多く、ある程度計画プロセスの標準化が進んだ分野では新規事業の事前評価は既存の知識や経験に基づき作成された事業計画の実行可能性を検証することが主眼となる。しかし、アフリカ諸国にとって持続可能な工業化やそのための企業家人材の育成は前例事業の少ない課題であり、その取組には企画案の作成段階から意図的に評価的思考を重視した事前評価が必要となる。本稿での事例報告は、Owen（2006）の「事業前評価（proactive evaluation）」のフレームワークを用いて、筆者が自ら関与した構想段階の新規企画に関して実施した事前評価についてであった。

事業前評価は、先行事例の少ない革新的事業を企画する際に極めて有用な計画ツールと考えられる。その実施は、事業構想の妥当性、実施可能性の検証から事業の新規企画の作成までを範囲とする。その有効性は、事業実施に関係するステークホルダーの積極的参加によって担保される。本稿で取り上げた国連大学ESDAプログラムのアフリカの企業家人材育成事業構想では、

事業の基本的な構図がアフリカの多数のビジネス大学院による連携協力を国連大学が誘導する形を取っており、予備的作業も含め事前評価を国連大学の事業担当チームが内部的に実施した。調査の過程で事業の最終受益者に当たる若手企業家との対話から始めてアフリカのビジネススクール訪問による聞き取り調査を実施し、さらに最終段階ではビジネス大学院の研究科長による検討のための会議を開催して、それら一連のデータ収集活動の分析結果を評価判断の根拠とした。事業前評価の有用性を担保するのは、ステークホルダーとの丁寧な対話からの信頼に足るデータの掘り起しであると考えられる。

一般的に事前評価の有用性を左右する要件はデータの存在である。本事例報告のように前例の少ない新分野ではデータが存在しないことの方が多い。特に統計データの欠落は顕著である。この課題に対応するには、聞き取り、フォーカス・グループ等の質的方法をうまく取り込むことが第一である。今回の調査ではワークショップを多用した。関連分野の異なるステークホルダーが集まり、それぞれの知識・経験を持ち寄ることで、情報の共有が容易にできるだけでなく、新たな知識を集合的に創出することも可能となる。

事業前評価の結果は、事業事務局を務める国連大学が当初想定していた参加ビジネス大学院による共通修士プログラムのような単一の企業家人材育成プログラムの創設に導くものではなかった。その代わりに複数の取組のパッケージをベースとして、ステークホルダーとの協働により企画案が確定し、資金的支援が得られた取組から実行に移す計画になった。具体的にはアフリカのビジネス大学院の間で相互連携の環境づくりを向こう5年をかけて実施することとした。そのためにアフリカの有力ビジネス大学院15校がそれぞれの志向に応じて参加する複数の課題別「実践コミュニティ（Community of Practice）」を形成するとともに、プログラム創設に向けての共同研究・研修の実行案づくりとドナーによる支援獲得のための作業を進めることになった。新型コロナの蔓延等もあって、事前評価の結果を活用した新規企画の具現化は遅れがちである

が、事前評価の過程で若手企業家から強い関心の表明のあったアジア・アフリカ企業家交流プログラムについては、2021年になってドナー二機関からの資金的支援が得られることとなり、事前評価の結果の活用が実現することになった。新規事業の創設には周辺状況の変化がつきものである。それに対応するには事前評価と取り組む側に柔軟な姿勢と方法的なレジリエンスが求められる。事業前評価の活用はその双方の重要性を教えてくれた。

謝辞

本事例報告の基になった研究は、JSPS科研費JP16K13348の助成を受けたものです。

本事例報告の本稿に対して、査読者の方々から有益なご指摘をいただきました。この場を借りて、深く感謝いたします。

注記

- 1 ESDA (Education for Sustainable Development in Africa) については、下記を参照されたい。 <https://unu.edu/projects/education-for-sustainable-development-in-africa->
- 2 世界銀行データによると、サハラ以南地域の実質タームの年平均経済成長率は、1980年代の2.1%、1990年代の2.3%からより長期の2000–2016年間に5.3%と顕著に加速した (World Development Indicators 2017年)。
- 3 Owen (2006) は、Plan-Do-Check-Actionの事業サイクルの進行にともなって発生する評価目的に対応する評価の異なる基本型を提示している。それによると、事前 (Plan段階) が事業前評価、期中 (Do段階) が事業形成評価 (Clarificative evaluation) とモニタリング評価 (Monitoring evaluation)、期末・事後 (Check段階) がインパクト評価 (Impact evaluation) で、事業サイクル上のどの段階でも使われる対話型評価 (interactive evaluation) を加えて5つの基本型から成る評価体系としている。
- 4 UNU-IAS (2016) . “Developing an Exploratory Research Programme on the Role of Youth Entrepreneurship for Sustainable Development in

- Africa,” ESDA Workshop Summary Report, Cape Town.
- 5 ビジネススクールは、学生の多くが社会人であること、財政的支援も含めて企業との直接的な結びつきが強いこと、大学から物理的に離れた市街地にキャンパスを構えていること等に起因する。
 - 6 会議の内容と結果の詳細については、UNU-IAS (2016) . “Report on African Youth Entrepreneurship Initiative,” Nairobi, Kenya.を参照されたい。
 - 7 比較対象とした6校は、南アフリカのケープタウン大学ビジネス大学院とプレトリア大学GIBSビジネススクール、ケニアのナイロビ大学ビジネス大学院とストラトモア・ビジネススクール (私立)、ガーナのガーナ大学ビジネス大学院、ナイジェリアのラゴス・ビジネススクール (私立)。
 - 8 ESDA Secretariat (2017) . *ESDA MBA Programs Comparison*, internal memorandum.

参考文献

- 国際協力機構 (2004) 『プロジェクト評価の実践的手法 JICA事業評価ガイドライン改訂版』、国際協力出版会
- Ács, Z.J. and Naudé, W. (2012) . *Entrepreneurship, stages of development, and industrialization*, MERIT Working Papers 021, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT) .
- African Development Bank (2015) . *At the Center of Africa's Transformation: Strategy for 2013–2022*, Tunis.
- African Development Bank, OECD and UNDP (2017) . *African Economic Outlook 2017: Entrepreneurship and Industrialization*, (<http://dx.doi.org/10.1787/aeo-2017-en>) .
- African Union Commission (2015) . *Agenda 2063: The Africa We Want*. Addis Ababa.
- Kapfudzaruwa, F., Nagao, M. & Mutisya, E. (eds) . (2018) . *Youth Entrepreneurship and Africa's Sustainable Industrialization*. Denver: Spears Media Press.
- Khan, A. R. (2011) . Employment in Sub-Saharan Africa: Lessons to be Learnt from the East Asian Experience. In Noman, A., Botchwey, K., Stein, H. and Stiglitz, J.E. (eds.) (2011) . *Good Growth and Governance in Africa: Rethinking Development Strategies*. Published to Oxford

- Scholarship Online: May 2012. DOI: 10.1093/acprof:oso/9780199698561.003.0015.
- Naude, W.A. (ed.) (2011) . *Entrepreneurship and Economic Development*. Basingstoke: Palgrave Macmillan.
- Newman, C., Page, J., Rand, J., Shimeles, A., Soderbom, M., and Tarp, F. (2016) . *Made in Africa: Learning to Compete in Industry*, Washington, D.C., Brookings Institutions Press.
- Noman, A., Botchwey, K., Stein, H. and Stiglitz, J.E. (eds.) (2011) . *Good Growth and Governance in Africa: Rethinking Development Strategies*. Published to Oxford Scholarship Online: May 2012. DOI: 10.1093/acprof:oso/9780199698561.001.0001.
- Ohno, I. and Ohno, K. (2011) . Dynamic Capacity Development: What Africa Can Learn from Industrial Policy Formulation in East Asia. In Noman, A., Botchwey, K., Stein, H. and Stiglitz, J.E. (eds.) (2011) . *Good Growth and Governance in Africa: Rethinking Development Strategies*. Published to Oxford Scholarship Online: May 2012. DOI: 10.1093/acprof:oso/9780199698561.001.0001.
- Owen, J. (2006) . *Program Evaluation, Forms and Approaches*, 3rd Edition, Crow's Nest, Australia: Allen & Unwin.
- Page, J. (2013) . Should Africa Industrialize? In Szirmai, A., Naudé, W. and Alcorta, L. (eds.) (2013) . *Pathways to Industrialization in the Twenty-First Century: New Challenges and Emerging Paradigms*. Published to Oxford Scholarship Online: May 2013. DOI:10.1093/acprof:oso/9780199667857.003.0009.
- Szirmai, A., Naudé, W. and Alcorta, L. (eds.) (2013) . *Pathways to Industrialization in the Twenty-First Century: New Challenges and Emerging Paradigms*. Published to Oxford Scholarship Online: May 2013. DOI: 10.1093/ acprof:oso/9780199667857.001.0001.

(2021.7.6 受理)

Evaluation Use in Framing New SDG Projects: Case Report of an African Youth Entrepreneur Capacity-building Project

Masafumi Nagao

International Development Center of Japan
Nagao169@gmail.com

Abstract

Ex-ante evaluation is usually conducted for the purpose of evaluating the contents and feasibility of a proposed plan for initiating a new project. However, there is observed increasing cases of SDGs-related projects which involve trans-disciplinary or multi-sectoral arrangements and which require delicate qualitative judgements on their outcome, necessitating ex-ante evaluation based on deliberate application of evaluative thinking from the planning stage. The present case report presents an ex-ante evaluation the author conducted which relied on J. Owen's framework of Proactive evaluation in order to combine planning and evaluation upfront. The two lessons learned from this case are the importance of practicing stakeholder dialogues from the initial stage of project conception and the possibility to effectively combine qualitative data collection methods such as interviews and focus groups with collective data-generating methods like workshops.

Keywords

Evaluation case report, Ex-ante evaluation, Proactive evaluation, African capacity-building,
Entrepreneurial capacity

【研究ノート】

プログラム評価の一類型としての 「社会的インパクト評価」の課題と可能性

伊藤 健

慶應義塾大学大学院政策・メディア研究科

特任講師

itok@sfc.keio.ac.jp

玉村 雅敏

慶應義塾大学総合政策学部

教授

tama@sfc.keio.ac.jp

植野 準太

慶應義塾大学大学院政策・メディア研究科

博士課程

uenojunta@keio.jp

要 約

社会的インパクト評価は、社会的企業や社会的投資の文脈で発展したプログラム評価の手法の一類型として捉えられる。社会的インパクト評価は、これまでのプログラム評価等の取り組みにおいて蓄積された知見を、非営利組織による社会的事業や、企業活動や事業投資に対して適用し、その社会的影響に対する評価の手法として活用しようとするものであるが、その定義やプログラム評価との関連性について十分な整理がされてこなかった。社会的インパクト評価がその主たる評価の対象とする社会的企業や社会的投資は近年大きく拡大しており、今後の評価実践の中で大きな部分となることが想定される。本稿では、これらの社会的インパクト評価の取り組みがどのように形成されたのか、プログラム評価を中心とした評価研究のなかで、どのように位置付けられるのか、関連する事象を整理したうえで、これらの評価の実践と理論的發展が、評価研究にどのように貢献しうるのかについて考察する。

キーワード

プログラム評価、社会的インパクト評価、社会的インパクト・マネジメント、
社会的投資、SDGs

1. はじめに

プログラム評価を「評価とは、プログラムや政策の実施プロセスもしくは成果を、明示的あるいは明示的な基準と比較しながら、体系的に明らかにすることであり、プログラムや政策の改善に貢献するための手段である」とした定義(Weiss 1995)によれば、近年社会的投資や非営利事業において「社会的インパクト評価」として呼ばれる一連の評価の取り組みは、プログラム評価の一類型として位置付けることができる。社会的インパクト評価は、これまでのプログラム評価や政策評価等の取り組みにおいて蓄積さ

れた知見を、社会的投資、非営利組織による社会的事業や、企業活動による社会的影響の評価に対して適用しようとするものである。

社会的インパクト評価に関する概念はSocial Impact Measurement (SIM)¹や Social Impact Assessment (SIA) 等の複数の概念から発祥している。特に、評価によって得られた知見を組織運営や企業経営、社会的投資の実践において活用しようという動きは、近年ではImpact Measurement and Management (IMM)²、あるいは日本語の実践的文脈においては「社会的インパクト・マネジメント」とも呼ばれているが、社会的インパクト評価、社会的インパクト・マネジメントと

いった複数の用語や概念についても、その定義やプログラム評価との関連性について十分な整理がされていないのが現状である。

また、こうした社会的インパクト評価に関する様々な活動は、学術的な研究よりも、社会的事業の実践において発展してきたという経緯もあり、必ずしも評価研究のなかで中心的な議題として取り上げられてこなかった。しかし、社会的インパクト評価の主な需要者であるESG投資、インパクト投資や社会的企業セクターに係る多様なアクターの発展に伴って、その手法の標準化やガイドラインの設定に対する議論が行われるようになった。

例えば、上場株式や債券に対する社会的投資であるESG投資は、2018年時点でグローバルには30兆ドル、日本において230兆円を超え（GSIA 2018）、狭義の社会的インパクト投資も2019年時点で3,179億円（上記は共に残高ベース）を超え（GSG国内諮問委員会 2019）ている。単純な比較は難しいが、2019年度の日本のODA予算が5,566億円であることを考慮すると、その社会的目的での資金の流れとしての存在感は明確に高まっており、こうした投資家が投資による社会的な成果を測るための手法に対する社会的需要が顕在化している。

特に社会的投資を実践する金融機関や投資家を中心に、グローバルに2,000以上のユーザーを持つImpact Management Project（IMP）が2016年に設立され、2008年から活動する民間の非営利組織であるGlobal Impact Investing Network（GIIN）が提唱する指標データベースが2019年にはIRIS+として更新、また2019年には国連開発計画により、SDGsに基づいた投資や事業の認証の仕組みとしてSDGs Impactの認証基準案が発表される等、国際機関や主要な機関投資家、ネットワーク組織等が国際的な評価基準の合意を議論するようになった。こうした動きに呼応する形で、特に企業や金融機関等においては、事業の非財務情報の開示の一環として、あるいはSDGs等によって提示された社会課題に対する事業による解決への貢献を可視化するために、社会的インパクト評価の手法を参照し、活用しようとする動きが定着を見せ始めている。

日本においても、社会的インパクト評価は、2015年に内閣府によって実施された「共助社会づくり懇談会」に設置されたワーキンググループにおいて、その定義が「社会的インパクト評価とは、社会的インパクトを定量的・定性的に把握し、当該事業や活動について価値判断を加えること」（内閣府 2016）とされ、この定義は、休眠預金基金事業に対する成果評価の手法の根拠の一つともなっている。

本稿では、これらの社会的インパクト評価の取り組みがどのように形成されたのかについて紹介し、これまでのプログラム評価を中心とした評価研究のなかで、どのように社会的インパクト評価が位置付けられるのか、関連する事象を整理したうえで、これらの評価の実践と理論的發展が評価研究にどのように貢献しうるかについて考察する。

なお、社会的インパクト評価という用語については、2（1）に後述する事情で、その対訳語や関連概念が明確にされないままに日本語での語彙として使用されているが、本稿がその用語の指す評価取り組みや関連する事業の定義やスコープを考察の対象とすること、またその語意に複数の異なるニュアンスがあることを前提に、Social Impact Measurementの対訳語として、また日本語での関連概念を指す用語として社会的インパクト評価として使用する。

2. 社会的インパクト評価の概念の背景と発祥

（1）社会的インパクト評価の概念と背景

社会的インパクト評価という語彙は、諸外国語において確立した対訳語が存在するわけではないが、関連する概念として、Social Impact Assessment（SIA）、あるいは Social Impact Measurement（SIM）を挙げることができる。

SIAは、例えば1992年にInterorganizational Committee on Guidelines and Principles for Social Impact Assessment（ICGP）が発表したガイドライン（ICGP 1992）にその評価の方法が定義されている。ガイドラインでは、環境インパクトアセス

メントの手法に対比する形で、環境以外の社会的要因に対する影響のアセスメントを指す概念としてSIAが定義された。SIMについては1990年代後半から社会的影響の把握を行う手法の総称として使用されるようになり、特に英米において90年代から先行して実践が行われたSocial Return on Investment (SROI、社会投資収益率)はしばしばSIMの代表的な事例として取り上げられている。SROIはその評価プロセスの中に、投下された費用と貨幣価値換算された社会的成果を比較する(The SROI Network 2012)ということからも、社会的活動の成果をその投資対効果の側面に着目して評価する、効率性評価としての性格の強い評価のあり方として認識できる。

プログラム評価におけるインパクト評価は、成果に対する介入の帰属性を重視し、「インパクト」とは介入によってもたらされたネットの効果(純効果)を指す(源 2020)が、社会的インパクト評価は、この定義におけるインパクト評価と同義ではないことに留意が必要である。また、国際開発の領域では、中長期的なアウトカムを「インパクト」と定義していることも、受け手によって社会的インパクト評価の異なる解釈を生じさせている。

日本における議論においては、社会的インパクト評価という用語が、公的機関で議論されたのは先述の内閣府におけるワーキンググループで議論されたのがその端緒であり、「社会的インパクト」は「短期、長期の変化を含め、当該事業や活動の成果として生じた社会的、環境的なアウトカム」とされた(内閣府 2016)ことから、日本語における社会的インパクト評価は、中長期的なアウトカムに対する成果の測定を行い、アウトカム評価や効率性評価を中心的な手法とする、プログラム評価の一形態として理解されていると言えるだろう。

(2) 「Social Impact Measurement」の対訳語としての「社会的インパクト評価」の妥当性

日本における社会的インパクト評価の取り組みが、社会的投資や非営利組織管理の文脈で議論された背景は、行政評価における業績測定についての議論と同様な問題意識が存在したと理

解できる。つまり、事業実施による成果を検証しようとする問題意識に基づく業績測定(performance measurement)の必要性に対する認識である。SIMという語法も、Social Impact、つまり事業のアウトカムを、定量的に評価可能な指標に基づいて測定(Measurement)を行うという語意であると理解でき、字義に基づいた日本語における訳語を充てるとすれば、「社会的インパクト測定」、あるいは行政評価の文脈で用いられる業績測定がそれにあたると考えられる。

しかしながら、後述する社会的インパクト評価の事例を参照すれば、社会的インパクト評価として定義される手法の多くは、プログラム評価におけるアウトカム評価、あるいは効率性評価の側面に着目しつつも、実際には、セオリー評価やプロセス評価を含む、プログラム評価の多様な側面を包含するものであり、ここに社会的インパクト測定ではなく、社会的インパクト評価という訳語が当てられた背景の一端があると考えられる。

その理由として、実際の社会的インパクト評価の導入においては、プログラム理論についての検討が、事業設計の一部として十分に行われてこなかったという課題に、多くの社会的インパクト評価が直面していることがある。評価の前提となる評価指標の設定のためには、事業目標や、事業のアウトカムとアウトプットの関連性の特定が必要であり、セオリー評価を行う必要があったことも、ロジックモデルの活用によるセオリー評価が、社会的インパクト評価の重要な手法として理解された理由の一つである。

行政評価の文脈での「業績測定」がプログラム評価と相互補完的な役割があるとすれば、社会的インパクト評価は、社会的投資の領域における「業績測定」であり、同様にプログラム評価との相互補完性を持つ評価の取り組みだと言える。業績測定も、その字義からいえば行政の政策実施による業績を定量的な手法で測定するものであるが、実際には単にその業績を明らかにするのみならず、業績評価で得られた情報をどのように政策実施のロジックに反映するかという評価デザインにも議論が及んでいる点においても、社会的インパクト評価の議論との相似

がある。

一つの事例として、日本における社会的インパクト評価の実践事例の一つである、休眠預金を活用した助成金運営を行う一般財団法人日本民間公益活動連携機構（JANPIA）の公開した社会的インパクト評価ガイドラインにおいても、社会的インパクト評価は、セオリー評価、アウトカム評価、インパクト評価、プロセス評価の複数の評価の性質を包含するものであるとしており、上記の認識を裏付けるものとなっている。

(3) 社会的インパクト評価の役割に対する国内外における評価

前述の通り、社会的インパクト評価が、社会的に、あるいは評価研究の発展にどのような役割を果たすのかについては、国内外で様々な議論がある。

馬場（2019）は、社会的インパクト評価について、特にその休眠預金基金の運用に対する適用についての課題を、英国での同様な評価の導入と比較する形で、(1) 評価の負担 (2) SROIとの混同 (3) 見せかけの成果（インパクト・ウォッシング）(4) 社会的インパクト評価では測れない価値の4点における課題を指摘しながら、「社会的インパクト評価は、多種多様な社会的価値を正確に測定するものではないが、多様な利害関係者が参画しながら、お互いに合意できるアウトカムの測定方法を探り、成果指標へと落とし込んでいくツールとして有用性を備えている」としている。

また、塚本（2020）は、英米におけるインパクト評価の潮流を振り返り、上述のインパクト評価の定義について触れた上で、日本における社会的インパクト評価が、インパクト評価の研究・実践や、これまでの評価研究の文脈を踏まえた議論になっていないこと、また社会的インパクト評価を謳って実践される評価の質について疑問を提示し、あたかも社会的インパクト評価が新しい枠組みであるかのような言説による弊害を指摘しながら、本来的なインパクト評価の推進のために、「反事実」を踏まえた評価の設計、標準化や質の向上のための官民連携が必要であると指摘している。

国外においても、同様な議論がある。例えば2016年の米国評価学会の年次大会では、米国評価学会とSocial Value Internationalとの共催で、サイドイベントとしてIMPCON2016が開催された。この開催に関して、Voは、American Evaluation Journal における寄稿論文において、評価（Evaluation）とインパクト評価（Impact Measurement, IM）¹は同じルーツを持つが、IMは特に社会的投資からのニーズへの対応として考案された手法であると位置づけている。評価とIMは、その異なる目的意識のために、評価のための資金源、得られた知見の使用者、評価アプローチ、評価単位や手法等の領域において違いがあると指摘し（表1）、この2つの領域の間での対話の重要性を提示している（Vo and Christine 2018）。

これらの議論は、2（2）で記述した、SIM（あるいはVoの用語ではIM）が、社会的投資における業績評価に該当し、相互に補完する取り組みであるという議論が、米国の文脈でも行われていることを示している。

3. 社会的インパクト評価枠組みの事例

社会的インパクト評価の枠組みには、前述のICGPIによるガイドラインやSROIのほかにも、国際機関や企業が提唱するものを含めると多数のフレームワークが発表されており、2011年時点でMaasは、30の異なる手法についての比較研究を行い、その類型について論じている（Maas 2011）。本節では、社会的インパクト評価の手法について、実践事例が多い事例としてSocial Value InternationalとImpact Management Projectの2者によるガイドラインを事例として取り上げ、その特徴と差異について述べる。

(1) Social Value International (SVI) によるガイドライン

SVIの前身はSROI Network Internationalであり、その設立の2008年から、1990年代に米国で開発されたSROIに代表される社会的インパクト評価の手法について研究を行い、欧州を含めた複数の地域における普及促進を行ってきた。SROIの

ガイドラインは2012年に発表された“A Guide to Social Return on Investment”までに複数回の更新が行われた。

SROIのガイドラインは、7つの原則と6つのステップから構成されている。そのプロセスは、ステップの1から3までは、分析のスコープとキーとなるステークホルダーの特定、アウトカムのマッピング、指標の設定とデータの収集となっている。ステップ4と5では、収集した指標データの貨幣価値換算、ステップ6では分析の報告と評価から得られた知見の組織への定着であり、特にステップ4と5は、アウトカムの貨幣価値換算での評価を謳うSROI特有なものである。

SROIが提唱したアウトカムの貨幣価値換算という概念は、特に当時黎明期であった社会的投資家や、一部の先進的な助成財団には歓迎されたが、貨幣価値換算に必要な係数が未整備であり、また特に事業による社会的インパクトの寄与率を算出する方法についてのガイドラインが不足していたため、評価基準の恣意性を指摘する声も多くあった。

SROI Networkは、2016年にSocial Impact Analyst Association (SIIA) と合併してSocial Value International (SVI) と改称し、その翌年の2017年には、“Maximize Your Impact – A Guide to Social Entrepreneurs”を発表している。

Maximize Your Impactでは、社会的インパクト評価を4つのステップに定義し、Plan, Do, Assess, Reviseの各プロセスを示しており、貨幣価値換算を必須のプロセスとしていない。また、10 Impact Questionsを設定し、4つのステップのそれぞれにおいて、何を明らかにするべきかについてのキー・クエスチョンとして定めることで、社会的インパクト評価の実践のガイドラインとしている。

この2つのガイドラインは、どちらもSVIが提唱する7つの原則³に則ったものであるとされ、2021年現在も、ウェブサイト上の掲載や研修の実施についてもSROIとMaximize Your Impactの2つのトレーニングが並行して実施されている。

(2) Impact Management Project (IMP) による ガイドライン

IMPは、2018年にUnited Nations Development Program (UNDP)、International Finance Corporation (IFC)、Organisation for Economic Co-operation and Development (OECD)、Global Reporting Initiative (GRI)、SVI、GIIN、World Benchmarking Alliance、GSGの8つの組織が共同で設立した取り組みである。IMPにおいては、社会的インパクトの評価の事業への導入は「社会的インパクト・マネジメント」と呼びならわされ、企業や投資機関の運営プロセスやマネジメント・サイクルに、社会的インパクトが組み込まれることがその本質的な価値実現の形として奨励されている。これは、これまでの事業評価が、単なるレポーティングの一環として実施され、得られた知見が事業へのインプットとして活用されないという課題意識に基づき、社会的インパクト評価の結果得られた情報が継続的に事業に提供され、経営のプロセスの一環となることで、事業の社会的価値が継続的に向上するという理論であると理解できる。

2018年に発表されたImpact Management Normsは、IMPの提唱する社会的インパクト・マネジメントのガイドラインとなるものである。Impact Management Normsは企業向けと投資家向けの2つの部分から構成され、共通する枠組みとして、社会的インパクトをWhat（どのような社会的変化が発生するか）、Who（受益者は誰か）、How Much（成果の規模、深度、持続期間はどれだけか）、Contribution（事業が存在しない状態に比較した純粋な貢献はどれだけか）、Risk（事業が期待通りに成果を実現しない場合のリスクは何か）の5つの観点からどのように捉えるかという5つの次元（Dimensions）、またそれに基づく評価の観点として15のカテゴリーの提示が行われている。

IMPのフレームワークでは、事業における社会的インパクトに対する取り組みが、5つの次元、あるいは15のカテゴリーそれぞれに、A（Avoid to harm）、B（Benefit Stakeholders）、C（Contribute to the solutions）の3つのレベルのどれにあたるかを評価し、特にHow Muchについてはアウトカムに定量的な指標を設定して、社会的インパクト

の評価を行うことが定められている。

(3) 2つのガイドラインが示す社会的インパクト評価の手法開発の趨勢

IMPの社会的インパクト評価への手法的貢献は、特に社会的インパクト評価の対象となる社会的変化の特定について、5つの次元の観点を提示、これまでSVIによるフレームワークにおいてはガイドラインの中で文脈として触れられていた要素を整理した点にある。また、ABCのカテゴリによって、どのようなレベルの社会的インパクトを志向するのか、あるいは実現できたのかについての評価基準を提示したところも、手法上の差分であると言える。

それに対して、社会課題の解決への貢献に至る道筋をセオリー評価として分析するという意味では、SVIのガイドラインでは、SROIにおけるインパクト・マップ、Maximize Your Impactにおけるセオリー・オブ・チェンジ（両者は、いわゆるロジックモデルに該当する）を組み立て、事業の社会的価値創造のプロセスを分析することを定めているのに対して、IMPにおいてはその部分について特段の記載が存在しない。

ロジックモデルが複数の社会的インパクト評価の手法の中で共通した要素として取り上げられる理由は、評価で得られた知見を提供するにあたって、事業のセオリーの理解が必要だからである。評価による事業へのフィードバックとして、事業が実現するアウトプット、あるいはアウトカムの評価の実施により、成果指標のデータを取得し、事業プロセスについて分析することで、社会課題の解決への貢献に寄与する知見やエビデンスを提供することへの期待がある。

ところが、IMPのフレームワークでは、事業のプロセスや、成果をもたらすセオリーについての情報は提供されないため、これについては別途事業主体や評価者が追加の評価作業を行う必要がある。投資的なアプローチでは、収益性の低い事業は持ち分を売却し、より高い収益を実現する事業と入れ替えることで、事業ポートフォリオ全体の収益性を保つことが一般的だが、社会的インパクトの観点でのポートフォリオ管理においても、個別事業の改善に踏み込まない

のであれば、こうした評価の手法が必要十分であり、却って費用対効果に優れた社会的インパクト評価の手法とされることも考えられる。

こうした特色を持つIMPのフレームワークは、社会的インパクト評価が、社会的投資に適合したプログラム評価のあり方の一類型であるというVo（2018）の指摘にも符号する。

4. 社会的インパクト評価の手法的特色と、評価実践に対する貢献

Rossiは、評価には（a）ニーズ評価（b）セオリー評価（c）プロセス評価（d）アウトカム／インパクト評価（e）効率性評価の5つの階層が存在すると指摘し、プログラムの設計、実施、介入後の評価のそれぞれの段階において異なるフォーカスがありうるとしている。

また、源（2020）は、このRossiによる定義を踏まえ、プログラム評価の要素として①社会的調査手法の活用②プログラムの有効性の体系的調査③基準との比較④政策的・組織的な文脈を考慮した実施⑤社会状況を改善するための活動に資する情報を提供するという5つを挙げ、プログラム評価が期待される社会的役割を果たすためには、どのような条件を備える必要があるかについて条件を提示している。

こうしたプログラム評価の定義に鑑みれば、社会的インパクト評価も、事業の実施プロセスや成果を明らかにし、社会状況を改善するための活動に資する情報を提供するという意味で、プログラムの評価の一類型であると理解することができる。一方で、社会的投資や社会的事業にその主なニーズが存在する社会的インパクト評価は、複数の異なる領域における評価実践への貢献をもたらしていることも事実である。

これまで概観したように、社会的インパクト評価は、社会的投資市場や社会的企業の伸長をその背景に、投資家や社会的企業に特有なニーズに応える形で、その手法を発展させてきた。

以下の節では、既存の評価理論の中で、社会的インパクト評価の特徴がどのように位置づけられるのか、またその特徴によって、評価理論

の発展に対してどのような貢献の可能性、あるいはリスクが存在するのかについて考察する。

(1) アウトカム／インパクト評価としての社会的インパクト評価の可能性とリスク

社会的インパクト評価の第一の特徴は、その評価の主眼が、プログラムの介入によって、介入の対象や社会全体に現れる変化としてのアウトカムや、中長期的なインパクトに着目した評価であるということである。SROIを開発したRoberts Enterprise Development Fund (REDF)の当時の事務局長であったJed Emersonは、「非営利組織への資金提供がエビデンスを伴った成果に基づくものであることが重要だ」という問題意識をもとに、SROIの評価枠組みの構築を推進したと述べている⁴。

特にSROIが開発された90年代の非営利事業における評価の状況を鑑みると、成果の発現が中長期にわたり、領域によってはその成果指標の設定が複雑になる等の背景から、アウトカムレベルの成果評価が一般的ではなく、事業評価の多くがアウトプットのレベルでの成果報告にとどまっていたことを考えると、社会的事業のアウトカムに着目し、その効率性を測る手法や指標への開発需要があったことが理解できる。

一方で、評価の手法として評価理論で言うところのインパクト評価を想定し、一定の精度を求めるのであれば、事業の実施と、事業実施の成果とされる事象の間の因果関係の特定や、事業の実施による成果に対する純粋な寄与の特定等の面で、その実施には専門的な知見と、一定の工数と費用が必要となることは大きな課題である。社会的投資や営利事業として実施される社会的事業に対する評価は、行政事業や国際協力事業のアカウンタビリティの履行の一環として実施される政策評価や国際開発における評価に比べ、事業あたりの予算も、また事業予算における評価コストも相対的に小さい場合が多い。事業評価の費用がこうした小規模な事業に対する投資収益や事業収益から支出されることを想定すると、評価コストの高さが、社会的投資の実施の阻害要因となってしまうリスクが発生するというジレンマがある。

こうした状況から、社会的インパクト評価の活用には、社会的事業の社会的成果を最大化し、社会課題解決に資する事業となる可能性を提示する反面、本来的なプログラム評価の多様な機能を実現できないリスクも存在する。

(2) 効率性評価としての社会的インパクト評価の可能性とリスク

社会的インパクト評価に特徴的なもう一つの観点は、事業による投資対効果、すなわち効率性に対する着目である。例えば前述のSROIの手法は、事業の成果としての社会的変化に着目したアウトカム評価、ロジックモデルを活用したセオリー評価、ステークホルダー参加による参加型評価⁵等、複数の側面があるが、その名称が表す通り、投入されたリソースに対する社会的価値の比率を算出する評価プロセスが6つのステップの中に定義されている(The SROI Network 2012)ことから、効率性評価としての側面を強く持っていることが理解できる。

SROIが開発された1990年代の米国のフィランソロピー・セクターは、伝統的な慈善財団に加えて、金融や情報通信等の新興産業によって資産を形成した経営者や、投資家としての背景のあるフィランソロピストが多く参入するようになった状況があった。こうした事業的、投資的な観点を持つフィランソロピストの観点からは、従来の非営利組織による事業が、効率性の観点から適切なマネジメントを実施する必要がある、そうした取り組みによってより大きな社会的成果が得られる可能性がある」と議論された(Grossman 2013)。こうした議論に基づいて、ベンチャー・フィランソロピーのように、非営利事業において、事業的な手法での経営管理を実施し、社会的成果を最大化しようという取り組みが発展したことは、非営利組織における評価実践に対する大きな貢献であった。

しかしながら、プログラム評価がその評価の対象とする社会的事業の価値には多様な側面が存在する。特に社会課題解決の主体としての市民社会という観点では、社会課題解決に資するがどのような過程によってその追求する価値や、解決の手法を決定し、事業が実施されるのか、

またその意思決定が広範なステークホルダーによる参加によって決定されたのかというプロセスや、こうした活動に対する市民参加そのものが、民主主義の実践という観点で重要な価値として受け止められる。こうした価値観からすれば、事業の社会課題解決に寄与する効率性が重視され、プロセスや意思決定の主体の構成に多くの関心が払われなければ、逆に社会の民主的な価値を毀損しかねないという可能性は、社会的インパクト評価の効率性評価としての側面に対するリスクとして認識されるべきであろう。

(3) 社会的インパクト評価の比較可能性の発展に対する可能性とリスク

社会的インパクト評価のもう一つの特徴は、複数の事業に対する社会的インパクト評価の実施の結果得られた情報を、比較可能な形で活用する、比較可能性についての着目である。

社会的投資家による評価に対するニーズの一部は、投資先の事業の社会的パフォーマンスを比較し、より大きな成果を実現する事業に資金を配分することで、投資による社会的成果を最大化しようという動機である。

例えば、GIINが提供する評価指標のデータベースは、もし行政がその開発を行えば莫大なコストのかかるインフラであり、これを民間の資金と知見で構築し、2,000を超える機関投資家がこれを使用して社会的投資を実行し、評価実践を行っているとするれば、これらは評価の知見への大きな貢献であり、社会的事業の成長により、社会的課題の解決への貢献をもたらしているだろう。こうした標準化された指標データベースや、異なる領域にまたがる社会的影響に対する比較可能性の担保への取り組みは、社会的インパクト評価の評価リソースに対する貢献であると言える。

反面、社会的インパクト評価の発展段階における初期、例えばSROI手法による貨幣価値換算のアプローチによる比較可能性への取り組みは、財務係数の設定についてのガイドラインやデータベースが十分に整備されていなかったこと、また反事実のコンセプトは提示されたものの、比較対象群との比較やRCT等による純粋なインパ

クトの提示の手法の実施可能性が十分に示されなかったことによって、評価手法そのものへの信憑性に疑問が呈されることになった。こうした事象は、比較可能性を追求する社会的インパクト評価の使用者に対して、十分に洗練した手法を提示できないリスクが露呈した結果とも言える。

5. 社会的インパクト評価とプログラム評価の整合性と課題

本節では、これまで概観した社会的インパクト評価の特徴を踏まえて、社会的インパクト評価の手法的な課題と、それらの課題を克服するために期待される取り組みについて述べる。

(1) 上位概念としての最終アウトカムとの整合性

社会的インパクト評価の課題の一つには、個別の評価の前提条件として、個別の事業成果としてのアウトカムと、上位目標としての最終アウトカムとの関連性が明確に定義されていないことが多いことが挙げられる。

山谷によれば、プログラム評価とそれ以外の評価活動を区別する観点は、「システムとロジックの2つの考え方、2つを反映したデザインの3つである」とされる（山谷 2012）。ここでのシステムとは、政策（policy）、プログラム、プロジェクトの3層構造に対応する体系であり、ロジックはプログラムを形成するインプット、アクティビティ、アウトプット、アウトカムのプロセスの論理である（山谷 2020）。プログラム評価は、評価によって得られた知見を、ロジック（プロセスの論理）と、システム（政策、プログラム、プロジェクトの体系）に反映させるに足る評価デザインを備えたものであることが求められる。

こうした考え方は、社会的インパクト評価にも受け継がれているが、社会的インパクト評価が対象とする民間の社会的事業においては、政策評価における上位目標である政策目標にあたる最終アウトカムが明確に定義されていないことがある。社会的インパクト評価の舞台である、社会的投資の対象である民間事業者の事業は、

往々にして地域レベルでの社会的課題の解決にその目的を置いたものとなり、事業の実施の結果、社会全体の課題解決にその取り組みがどのような影響を果たすかというロジックが設計されているとは限らない。

これは、社会的投資が対象とする事業が不完全だということではなく、そもそも行政事業との比較においては小規模に実施される社会的事業は、期待される社会的変化は相対的に小さなものととどまる場合もあること、また事業実施の期間にも制約があること等から、中期的な目標は設定できても、長期的なアウトカムについては明確な目標として設定しにくいという背景がある。こうした状況が、社会的インパクト評価の実施目的についての混乱を招く課題となり、また評価の結果として期待される価値判断の基準が不明瞭となる混乱をもたらしている事象が見受けられる。

たとえ上位概念が明確でない場合にも、評価のプロセスでそれを明らかにすることは、評価の意義と捉えることもできるが、社会的インパクト評価の実施にオーナーシップを持つ投資家等の資金提供者が、そうした上位概念で定義されるべき社会的成果についての妥当性を明らかにできない場合には、当然ながら評価の基準が曖昧になり、本来的な役割を果たせないという状況に陥るリスクがある。

こうした状況を回避し、社会的インパクト評価がその機能を十分に果たすためには、社会的投資家、あるいは社会的事業を実施する事業者が、社会レベルの実現目標としての最終アウトカムや明確な事業ビジョンを持つことが期待される。

(2) 市場メカニズムに基づく社会課題解決アプローチの功罪

また、社会的インパクト評価という概念や用語が、市場メカニズムに基づいて社会課題を解決しようという社会的投資や社会的企業の文脈から生まれたことから、その評価のアプローチにおいてもボトムアップによる社会課題解決への貢献への志向が見られる。

政策評価では、評価の結果蓄積されるデータ

や知見が、将来的な政策の精度を上げ、有効な政策を立案・実行するためのエビデンスとして機能することが期待される。こうした働きかけの方法は、トップダウン型のアプローチであるとすることができる。

それに対して、社会的インパクト評価がその背景とする金融市場のメカニズムは、もちろん市場の効率性や透明性といった市場の質的側面にも依存するが、個別の事業やファンドのパフォーマンスは、その事業や投資手法、投資活動に帰属するものとして捉えられる。市場の発展を促すためには、まずは事業の採算性や潜在性が立証され、そうした事業の効率性が市場原理を通じてほかの事業に波及するというのが、ボトムアップのアプローチである。

こうした社会的投資、社会的事業の特性は、社会的インパクト評価は政策形成のエビデンスを提供することが第一義的な目的とされるのではなく、個別の事業に対して適切な評価を行うこと、またその社会的価値を明らかにすることによって、市場に対して実証事例を提供することがその目的意識の主眼に置かれることになる。

このトップダウンとボトムアップのアプローチは、両者ともに一定の精度を持った社会的価値の判断基準を作るという意味では共通しているが、制度・政策を司る公的セクターに対するフィードバックに資するエビデンスの構築への志向と、個別事業の評価、あるいは市場メカニズムを形成する事例の構築では、そこに要求されるデータや分析の精度は大きく異なることは容易に想像できるところだろう。

このような背景により、社会的インパクト評価は、社会的事業に提供される資金の市場拡大に貢献するが、往々にして、評価の実施によって得られるエビデンスとしての知見は限定的なものになる。具体的には、個別の事業の改善や、特定の投資家にとっての投資判断の基準を提供することにはなっても、政策評価がその目的の一つとする、政策形成のためのエビデンスの構築に資する知見やモデルを提供するといった役割を果たすことが難しい側面が存在する。

6. 社会的インパクト評価が市場と経済システムにもたらす変革の可能性と課題

これまで述べた通り、社会的インパクト評価は、社会的投資の領域における業績測定を取り組みだと言える。社会的インパクト評価は、その実施により、アウトカム評価、および効率性評価の社会的事業への適用可能性を提示することで、増大する民間セクターからの評価ニーズに応え、評価実践の量的拡大をもたらした。

社会的投資や企業の社会性評価の対象となる事業の市場規模は、将来的に行政事業や国際協力のそれを大きく上回ることが想定され、評価により蓄積される知見による社会的影響も、大きな力量となることが考えられる。特に、社会的投資に対する社会的インパクト評価の導入は、金融市場そのものの質を変化させる可能性があり、SDGsによって示された社会的課題の解決に必要なファイナンス・ギャップの解消には不可欠な要素である。

そうした期待に応えるためには、社会的インパクト評価を実践する投資家や事業者のリテラシーの向上、同時に社会的投資の文脈に適合した評価手法の発展が期待される。

また、社会的なインフラとしての評価の仕組みの構築という意味では、IMPやGSGが取り組んでいるように、各国政府、国際機関、国際的な非営利組織、機関投資家等のネットワークを構築して、国際的なコンセンサスを創出し、評価基準についての国際合意を構築してゆくことが期待される。

こうした社会的インパクト評価の普及促進、国際的な枠組み作りの取り組み、具体的な事例の創出等、異なるレベルでの取り組みが有機的に結合することによって、市場経済が持つ構造的な課題に対して、制度レベルでの変革をもたらし、社会課題の解決に寄与する可能性がある。

7. おわりに ―社会的インパクト評価が評価研究にもたらす貢献

本稿では、社会的インパクト評価を、社会的投資や非営利組織管理の領域での業績測定の手法として位置づけ、プログラム評価との関連性を明らかにし、また同時にその手法的特色について明らかにすることを試みた。行政評価における業績測定が、民間事業における業績評価の手法を行政事業に取り入れることで、その効率性や革新的なサービスの導入を推進する推進力としようとしたのと同様に、社会的インパクト評価は、行政評価や非営利組織評価、民間企業の経営手法において培われた業績測定の枠組みを、社会的事業の領域に導入しようとする取り組みである。特定の領域で開発され、発展した枠組みを、異なる領域へ適用することによる手法的発展は、評価研究全体に便益をもたらすものだ。

また、社会的インパクト評価の初期に開発されたSROIが、社会的事業における多様なセクターの協働を前提に、評価プロセスへのステークホルダーの参加を原則として位置づけたように、民間セクターが社会的事業を推進する場合には、行政等単一のセクターの事業実施と異なり、その持続可能性と社会的インパクトを両立するために、行政や非営利組織、あるいは受益者を含む地域コミュニティとの協働が必要となる。

こうした意味では、社会的インパクト評価は、業績測定の枠組みの民間セクターへの適用という側面のみならず、多くのステークホルダーが協働する参加型評価の新しい領域としての可能性について検討することができ、評価研究に対する貢献の一つとして、今後の研究課題と考えることができるだろう。

注記

- 1 Social Impact Measurement (SIM) は英文においてはしばしばImpact Measurement (IM) と略されて表記される。Vo (2018) におけるIMも、本稿で取り上げている「社会的インパクト評価 (SIM)」と同義である。なお、本文で記載した通り、このIMも、

- 日本語における「インパクト評価」も「社会的インパクト評価」を省略した語法であり、英語での Impact Evaluationを指すものではない。
- 2 Impact Measurement and Managementは、社会的インパクト評価を特に社会的投資のマネジメントに適用する文脈から生まれた用語であり、特にGIIN等の社会的投資を推進する組織や金融機関等において使用される。
 - 3 SVIにおける社会的価値評価では、1.ステークホルダーの関与、2.変化に対する理解、3.重要な物事を価値づける、4.マテリアリティのあるもののみを評価の対象とする、5.過剰な主張をしない、6.透明性の担保、7.結果の検証を行う、と7つの原則が定められている。
 - 4 社会的インパクトマネジメントイニシアチブシンポジウムにおけるインタビュー（2021年1月26日）
 - 5 SROIが掲げる7つの原則の第一は「ステークホルダーの関与」であり、社会的価値は多様なステークホルダーによって定義されることの必要性を謳っている。
- ### 参考文献
- GSG国内諮問委員会（2020）『日本におけるインパクト投資の現状（2020年12月24日修正版）』、GSG国内諮問委員会
- 伊藤健・玉村雅敏（2015）「社会的投資収益率（SROI）法の発展過程と手法的特徴」、『日本評価研究』、15（1）:41-55
- 小関隆志・馬場英朗（2016）「インパクト評価の概念的整理とSROIの意義」、『ノンプロフィット・レビュー』、16（1）:5-14
- 佐々木亮（2010）『評価論理』、多賀出版
- 塚本一郎・関正雄（2020）『インパクト評価と社会イノベーション』、第一法規
- 内閣府（2016）「社会的インパクト評価の推進に向けて」
- 馬場英朗（2019）「休眠預金活用における社会的インパクト評価の論点整理：イギリスにおけるインパクト評価との相違点」、『関西学院大学商学論集』、64（3）:43-54
- 山谷清志（2012）『政策評価』、ミネルヴァ書房
- 源由理子・大島巖（2020）『プログラム評価ハンドブック』、晃洋書房
- Dufour, B.(2019). Social impact measurement: What can impact investment practices and the policy evaluation paradigm learn from each other?, *Research in International Business and Finance*, Volume 47, 18-30.
- Epstein, M. J. (2015). Measuring and Improving Social Impacts: A Guide for Nonprofits, Companies, and Impact, Berrett - Koehler (「社会的インパクトとは何か」英治出版)
- Global Sustainable Investment Alliance (2018). 2018 Global Sustainable Investment Review
- Grossman A., Appleby S., and Reimers, C.(2013). Venture Philanthropy: Its Evolution and Its Future, Harvard Business School Case Study N9-3 1 3-111.
- Interorganizational Committee on Guidelines and Principles for Social Impact Assessment (1992). Principles and guidelines for social impact assessment in the USA.
- Mass, K. and Liket, K. (2011). *Social Impact Measurement: classification of methods*, Environmental Management Accounting and Supply Chain Management, Rotterdam, Erasmus University, 171-202.
- Olsen, S., and Ruff, K. (2018). The Need for Analysts in Social Impact Measurement: How Evaluators Can Help. *American Journal of Evaluation*, 39(3), 402-407.
- OECD(2015). Policy Brief on Social Impact Measurement for Social Enterprises.
- Rossi P, Lipsey M., and Freeman H. (2004). Evaluation: A Systemic Approach 7th Edition（大島巖他訳(2005)『プログラム評価の理論と方法』、日本評論社）
- Social Value International (2017). Maximize Your Impact – A Guide to Social Entrepreneurs.
- The SROI Network (2012). A Guide to Social Return on Investment.
- Vo, A. T., and Christie, C. A. (2018). Where Impact Measurement Meets Evaluation: Tensions, Challenges, *American Journal of Evaluation*, 39(3), 383-388.
- Weiss Carol (1997). Evaluation 2nd edition.
- Prentice Hall (佐々木亮 監修(2014)『入門 評価学:政策・プログラム研究の方法』、日本評論社)
- How do enterprises use impact performance data to set goals?, <https://impactmanagementproject.com/impact-management/how-enterprises-manage-impact/>, Impact Management Project, 2021年3月15日参照

表1 EvaluationとImpact Measurementの収束と発散の領域

	Evaluation	Impact Measurement
社会的ドライバー	説明責任 意志決定： - プログラムの改善 - サイクル終了時の評価	意志決定： - 財政の開発 - 持続可能性 - 戦略的計画
人的ドライバー	公的資金提供者（例：行政） 民間資金提供者（例：財団、企業、エンジェル／インパクト投資家） プログラムスタッフ プログラム裨益者	民間資金提供者（例：財団、企業、エンジェル／インパクト投資家） 社会起業家 公的資金提供者（例：政府）
最終利用者	公的資金提供者（例：政府） 民間資金提供者（例：財団、企業、エンジェル／インパクト投資家） プログラムスタッフ プログラム裨益者	民間資金提供者（例：財団、企業、エンジェル／インパクト投資家） 社会的起業家 公的資金提供者（例：政府）
評価アプローチ	文脈、投入、過程、生産 (CIPP) モデル 複雑性理論 クリティカルシステムヒューリスティック 発展的評価 エンパワメント評価 評価可能性アセスメント 参加型評価 理論に依拠した評価	社会価値会計 定量評価 (Value)
関心対象となるユニット	投入 活動 結果 成果	活動 結果
分析手法 / デザイン	事例研究 / 記述的デザイン 貢献度分析 実験的デザイン 混合法 (Mixed Method) モスト・シグニフィカント・チェンジ (MSC) 手法 アウトカムハーベスティング 事前事後での成果測定 準実験的デザイン	バランススコアカード 事例比較 ロジックモデル分析 モニタリング パフォーマンス測定 事前事後でのアウトカム測定 社会的投資収益率 (SROI)
得られた知見の性質	特定のシステム 特定の組織 特定のプログラム 特定のグループ 特定の場所	特定のプログラム 特定のグループ 特定の場所
普及の実践	実施中 サイクルの終了時	実施中 サイクルの終了時

(出所) Vo. A. T. (2018) p.386を筆者翻訳。なお、注記¹に記した通り、ここでのImpact Measurementは、Social Impact Measurementと同義の用語とする。

(2021.7.13 受理)

Social Impact Measurement as a Type of Program Evaluation — Challenges and Opportunities —

Ken Ito

Project Assistant Professor,
Graduate School of Media and
Governance, Keio University
itok@sfc.keio.ac.jp

Masatoshi Tamamura

Professor, Faculty of Policy
Management, Keio University
tama@sfc.keio.ac.jp

Junta Ueno

Graduate School of Media and
Governance, Keio University
uenojunta@keio.jp

Abstract

Social Impact Measurement (SIM) can be regarded as a type of program evaluation method that has been developed in the context of social business and social investment. SIM is an attempt to apply the knowledge accumulated in the program evaluation to social enterprises, corporate activities, and investment activities to use it as a method to evaluate their social impact. However, the definition and relevance of SIM to program evaluation have not been sufficiently clarified. In recent years, the market size of social enterprises and social investments has been expanding significantly, and it is expected to become a major part of evaluation practice in the future. In this paper, we will examine how these SIM efforts were formed, how they are positioned in the theories of program evaluation by summarizing the relevant facts, and discuss how the practice and theoretical development of SIM can contribute to evaluation studies.

Keywords

Program Evaluation, Social Impact Measurement, Social Impact Management, Social Investment, SDGs

【研究ノート】

留学生 30 万人計画の成果と課題 ー成長戦略、大学のグローバル化及び日本語教育との 関係からの考察ー

佐藤 由利子

東京工業大学

yusato@tse.ens.titech.ac.jp

要 約

留学生30万人計画は2019年に30万人の目標を達成したが、偽装留学生を増加したという批判も見られる。このため本稿では、上位計画である成長戦略や、大学等のグローバル化など主な方策にかかる政策文書や予算書の分析から、計画の政策的枠組みを確認し、その成果と課題を検証した。

分析の結果、計画の主要な目的は「優秀な留学生の増加」と「留学生の国内就職」と特定され、その成果指標を分析したところ、大学等のグローバル化は英語コースを増加し、直接入学者の増加をもたらしたが、入学時点で日本語力を求めないという利点が日本就職促進上の課題となっていることが判明した。

また、方策に入っていなかった日本語教育機関が、人数目標達成に大きく貢献したものの、優秀な留学生の増加は十分に達成されておらず、日本語教育機関に対する支援とモニタリングを拡充する必要性が示唆された。

キーワード

外国人留学生、大学のグローバル化、高等教育の国際化、留学生の就職、日本語教育機関

1. はじめに

2008年に開始された「留学生30万人計画」は「高度人材受入れとの連携」を謳い、それまでの留学生を育て、母国に帰す政策から、日本社会に引き留める政策へと転換する契機となった。2019年に外国人留学生数は312,214人に達し、留学生30万人の目標値は達成された。他方、ベトナム、ネパールなど非漢字圏からの留学生が増加し、留学生30万人計画が「偽装留学生」の流入を引き起こしたという批判もある¹。このため、留学生30万人計画の成果と課題を検証し、課題

の要因を解明することが必要である。

留学生受入れ政策の評価については、総務省（2005）が「留学生10万人計画」の評価を行い、佐藤（2010）は、2005年以前に日本留学した者の調査などから評価を行っている。また、谷口ら（2011）は、国費外国人留学生制度の成果・効果を調べ、総務省（2017）は、グローバル人材の育成に関する政策評価の中で、大学の国際化の取り組みと外国人留学生の受入れ状況を調査し、留学生の増加の大半は日本語教育機関と専門学校入学者によると指摘し、総務省（2019）では、高度外国人材の受入れに関する政策評価

の中で、留学生の日本就職の課題として、日本語能力の不足や日本特有の就職活動知識の不足などを指摘している。この他、工藤ら（2014）は、日本の大学の国際化と留学生政策をレビューして提言を行い、白石（2019）は留学生10万人計画から30万人達成までの歴史を概括し、二子石（2021）は留学生の入学経路と進学進路の分析から、計画の実態解明を試みている。さらに「留学生30万人計画」関係省庁会議（2021）は、30万人計画を検証し、大学留学生の伸びが日本語教育機関で学ぶ者ほど大きくないものの、彼らの8割が高等教育機関に進学することから、高度人材の育成プロセスにあり、国内就職者の割合も増加して、高度人材の国内定着が図られつつあるとしている。しかし、上位政策である成長戦略との関係性、また、5つの方策、特に相互補完性が想定されていた大学等のグローバル化との関係性を把握した上で評価した研究は、管見の限り見当たらない。

このため、本研究では、留学生30万人計画、成長戦略及び大学等のグローバル化にかかる政策文書や予算書の分析から、上位政策、関連政策、方策に至る同計画の政策的枠組みを特定し、これに基づき作成した評価の視点から、30万人の成果と課題、また、課題が生じた要因について総合的に考察する。

日本における政策評価の課題について、西出（2020:173-174）は、政策の決定権が立法機関にありながら、行政機関が政策効果に関する結果責任を問われ、政策検証システムを牽引せざるを得ないジレンマを分析し、山谷（2021）は、政策評価が、政策官庁を自負する官庁以外、実質的にプロジェクト評価になっており、「政策体系」を意識した評価が少ないことを指摘する。

本稿の意義は、留学生30万人計画について、上位政策や方策との関連性を考慮した政策評価を行うことにより、政策体系に基づく評価の事例を示すこと、また、第三者の立場で、同計画の成果と課題を明らかにすることにある。

2. 留学生30万人計画の政策的枠組み

（1）留学生30万人計画の背景と趣旨

福田康夫首相（当時）は2008年1月の施政方針演説において「新たに日本への留学生30万人計画を策定し、実施に移すとともに、産学官連携による海外の優秀な人材の大学院・企業への受入れの拡大を進める」と表明した²。この背景には、IT分野などで高度外国人材を積極的に受け入れる気運が高まったこと³、経済のグローバル化などに対応して、企業の海外展開が加速し、留学生の雇用促進の動きが強まったことなどが挙げられる。例えば、日本経済団体連合会（2004）は『外国人受け入れ問題に関する提言』の中で、優秀な外国人材を獲得する方策の1つとして、留学生の国内就職の促進を提言している。

2008年7月には、文部科学省（以下、文科省）、外務省、法務省、厚生労働省、経済産業省（以下、経産省）、国土交通省の6省連名で「留学生30万人計画の骨子」が発表された⁴。骨子は、趣旨と方策の2部構成で、表1は、趣旨の全文と方策の項目を示している。

趣旨の①から、「留学生30万人受入れ」自体が計画の目標であり、「高度人材受入れとの連携」、「国・地域・分野などに留意した優秀な留学生の戦略的獲得」が重視されていることがわかる。知的国際貢献は最後に記され、栖原（2010:9）は、従来、留学生受入れの主目的であった人材養成等の知的国際貢献は後方に押しやられ、高度人材受入れと優秀な留学生獲得が主目的になったと分析する。

趣旨②には「入口から卒業・修了後の進路に至るまで体系的に」「関係省庁・機関等が総合的・有機的に連携」と推進方法が記されている。

また方策には、留学生30万人達成に向けて、日本留学の動機付けと入学プロセスの円滑化（1と2）、在学中の教育と環境整備（3と4）、卒業・修了後の日本社会への受入れ推進（5）に関する具体的手段が列挙されている。

このように「留学生30万人計画の骨子」には、「留学生30万人の受入れ」の達成目標や方法は書かれているが、留学生30万人受入れを通じて達成する目的については明示されていない。

このため、文科省が、骨子発表の3ヶ月前に作成した「『留学生30万人計画の骨子』とりまとめ」の考え方を確認すると、「留学生30万人計画の意義」として、「我が国にとっての意義」が7項目、「我が国の大学等における意義」が7項目記載されている⁵（表2）。

このことから、留学生30万人計画とりまとめの中心であった文科省では、留学生受入れを通じ、国と大学等に多くの「意義＝成果」を想定していたことがわかる。特に、我が国にとっての意義の最初に掲げられた「我が国の科学技術、

産業等の国際競争力の維持・向上」と「我が国の経済活動の担い手として、労働市場に（優秀な）人材を確保」は、政府開発援助の一環として留学生の帰国を促してきた従来の留学生政策（佐藤 2010）に示されてこなかった項目である。

また、大学等における留学交流の意義には、大学等の国際的な通用性、国際競争力の強化や日本人学生の国際理解増進が挙げられている。30万人計画の方策3には、大学等のグローバル化の推進が、計画達成の手段として掲げられており、大学等のグローバル化と30万人計画には、

表1 「留学生30万人計画」骨子に記載された趣旨と方策

趣旨
①日本を世界により開かれた国とし、アジア、世界との間のヒト、モノ、カネ、情報の流れを拡大する「グローバル戦略」を展開する一環として、2020年を目途に留学生受入れ30万人を目指す。その際、高度人材受入れとも連携させながら、国・地域・分野などに留意しつつ、優秀な留学生を戦略的に獲得していく。また、引き続き、アジアをはじめとした諸外国に対する知的国際貢献等を果たすことにも努めていく。
②このため、我が国への留学についての関心呼び起こす動機づけから、入試・入学・入国の入り口から大学等や社会での受入れ、就職など卒業・修了後の進路に至るまで、体系的に以下の方策を実施し、関係省庁・機関等が総合的・有機的に連携して計画を推進する。
方策
1 日本留学への誘い～日本留学の動機づけとワンストップサービスの展開～
2 入試・入学・入国の入り口の改善～日本留学の円滑化～
3 大学等のグローバル化の推進～魅力ある大学づくり～
4 受入れ環境づくり～安心して勉学に専念できる環境への取組～
5 卒業・修了後の社会の受入れの推進～社会のグローバル化～

（出所）文部科学省、外務省、法務省、厚生労働省、経済産業省、国土交通省（2008）「『留学生30万人計画』骨子」

表2 「留学生30万人計画」の意義

1) 我が国にとっての留学生交流の意義
・我が国の科学技術、産業等の国際競争力の維持・向上
・我が国の経済活動の担い手として、労働市場に（優秀な）人材を確保
・発展途上国出身の学生を受入れ、高等教育を受ける機会を提供し、人材育成を通じた国際貢献を実施
・我が国と留学生の出身国・地域との国際親善の強化
・我が国と諸外国との間の人的ネットワークの形成により、相互理解と友好関係が深化し、世界の安定と平和に資する
・留学生を通じた日本語、日本文化等「日本」の魅力の普及、海外における我が国の理解者、支援者の育成
・留学生という若者の活力が少子高齢化を迎えた我が国又は地域を活性化
2) 我が国の大学等における留学生交流の意義
・大学等（大学、大学院、短期大学、高等専門学校、専門学校）の国際的な通用性・共通性を向上し、国際競争力を強化
・大学等の教育研究や国際的評価の向上
・知的国際貢献の実現
・国際的学術ネットワークの確立・進展
・少子化に対応した経営安定化
・日本人学生の国際理解増進や異文化体験、語学力向上
・キャンパスの国際的な環境の創出

（出所）文部科学省（2008）「『留学生30万人計画の骨子』とりまとめ」の考え方」

互いに目的と手段という、相互補完関係が想定されていたことがわかる。

(2) 留学生30万人計画の施策・事業と予算

2009年に内閣府と30万人計画関係6省が計画の進捗を取りまとめた資料⁶には、5つの方策別に施策・事業と予算が示されている（表3）。

予算（内数分を除く）が多いのは方策4「受け入れ環境づくり」、方策3「大学等のグローバル化の推進」、方策5「卒業・修了後の社会の受入れの推進」の順で、方策3では「国際化拠点整備事業（グローバル30）」に新規予算40.8億円が計上されている。2008年の教育振興基本計画において、留学交流の推進や国際活動の充実を通じた大学等の国際化が掲げられたことが背景にある。

日本人学生の海外留学推進予算が含まれるのは、産業界からのグローバル人材育成の要請が高まり（吉田 2014: 30-32）、政府が日本人の海外留学や双方向の学生交流の必要性を強く認識したこと（工藤ほか 2014: 35）が背景にある。

府省別では、文科省予算が全体の80.2%を占め、外務省の8.8%、経産省の6.8%が続く。30万人計画は6省が推進していたが、実際には文科省が中心的役割を果たしていたことがわかる。

(3) 成長戦略における記述分析

2013年、日本経済再生本部は「日本再興戦略」を発表し、戦略はその後、アベノミクスの第3の矢「成長戦略」として毎年更新され、国の最重要指針として機能してきた。

表4は、2013年～19年の成長戦略における30万人計画と大学等のグローバル化に関する記述を抜き出したものである。吉田（2012）は2000年代以降、グローバル化の名の下、経済界、経産省、文科省が高等教育政策に関して同一歩調をとるようになったと指摘するが、30万人計画と大学等のグローバル化も、成長戦略の中で経済成長の手段として位置づけられていることがわかる。

「日本再興戦略」（2013）の「一．日本産業再

表3 留学生30万人計画の主な施策・事業と予算（2009年度）

1 日本留学への誘い 20 億円
ワンストップサービスの展開
・日本留学ポータルサイト（文部科学省、新規 0.3 億円）、海外拠点等における相談体制（文部科学省、3.1 億円）
・留学生アドバイザーの配置による日本留学の各種相談（外務省、0.8 億円）
・ビジット・ジャパン・アップグレード・プロジェクト（国土交通省、33.3 億円の内数）、青年の船等（内閣府、15.8 億円）
海外における日本語教育の普及：（外務省・国際交流基金、125.7 億円の内数）
2 入試・入学・入国の入り口の改善 3.4 億円
留学生の受入れ拡大に伴う審査体制の充実・強化（法務省、26.3 億円の内数）
日本留学試験の拡充（文部科学省、3.4 億円）
3 大学等のグローバル化の推進 47.1 億円
グローバル 30 拠点の形成（文部科学省、新規 40.8 億円）
日本人学生の海外留学の推進（文部科学省、6.3 億円）
4 受入れ環境づくり 404.1 億円
留学生宿舍の確保
・大学等の留学生宿舍借上げ支援等（文部科学省、12.7 億円）
・あんしん賃貸支援事業（国土交通省、1.7 億円）
外国人留学生奨学金制度の充実
・国費留学生（文部科学省、220.2 億円）、私費留学生学習奨励費（文部科学省、78.9 億円）
・留学生交流支援（短期受入れ）（文部科学省、16.4 億円）、授業料減免学校法人援助（文部科学省、31.4 億円）
・人材育成研究支援無償（外務省、42.0 億円）、国連大学私費留学生育英資金貸与事業（外務省、0.8 億円）
5 卒業・修了後の社会の受入れの推進 43.2 億円
留学生の就職支援の充実
アジア人財資金構想（文部科学省、国費 220.2 億円の内数、経済産業省、34.0 億円）、現地産業人材（経済産業省、1.6 億円）
留学生のための就職情報提供事業（文部科学省、0.4 億円）、専修学校留学生総合支援プラン（文部科学省、1.3 億円）
外国人雇用サービスセンター等（厚生労働省、3.5 億円）、企業側の意識改革や受入れ体制整備（厚生労働省、0.5 億円）
帰国留学生会への支援の充実（外務省、1.9 億円）

（注）補正予算を除く

（出所）内閣府ほか（2009）「留学生30万人計画の進捗状況について」19-20

表4 成長戦略における「留学生30万人計画」と「大学等のグローバル化」に関する記述

	優秀な外国人留学生の増加	外国人留学生の国内就職促進	大学等のグローバル化
日本再興戦略 2013	一. 日本産業再興プラン 2. 雇用制度改革・人材力の強化 ⑦グローバル化に対応する人材力の強化 日本人留学生を6万人から2020年に12万人へ倍増 優秀な外国人留学生を2012年の14万人から2020年に30万人に倍増（「留学生30万人計画」の実現）	三. 国際展開戦略 3. 我が国の成長を支える資金・人材等に関する基盤の整備 ②潜在力ある中堅・中小企業等に対する重点的支援 外国留学生の就職あっせん支援（年間1万人の外国人留学生の我が国での就職を目指す）	一. 日本産業再興プラン 2. 雇用制度改革・人材力の強化 ⑥大学改革 大胆で先駆的な改革を後押し。 世界と競う「スーパーグローバル大学」を創設、今後10年間で世界大学ランキングトップ100に我が国の大学が10校以上入ることを目指す
日本再興戦略 改訂 2014	一. 日本産業再興プラン 2. 雇用制度改革・人材力の強化		
	2-3. 大学改革／グローバル化等に対応する人材力の強化 KPI：2020年までに日本人留学生を6万人から12万人へ倍増、2020年までに外国人留学生を倍増（「留学生30万人計画」の実現）	2-2. 女性／若者・高齢者等の活躍推進／外国人材の活用 KPI：ポイント制の導入後11か月で高度人材認定された外国人の増加 ①高度外国人材受入環境の整備「留学生の国内企業（特に中小企業）への就職拡大」	2-3. 大学改革／グローバル化等に対応する人材力の強化 KPI：今後10年間で世界大学ランキングトップ100に10校以上入る
日本再興戦略 改訂 2015	一. 日本産業再興プラン		
	2. 雇用制度改革・人材力の強化 2-1. グローバル化等に対応する人材力の強化 KPI：2020年までに日本人留学生を6万人から12万人へ倍増、2020年までに外国人留学生を倍増（「留学生30万人計画」の実現）	2. 雇用制度改革・人材力の強化 2-2 女性の活躍推進／外国人材の活用 KPI：2017年末までに5,000人の高度人材認定を目指す ii) 外国人材の活用 ②留学生の更なる受入れ加速化と留学後の活躍支援強化	3. 大学改革／科学技術イノベーションの推進／世界最高の知財立国 KPI：今後10年間で世界大学ランキングトップ100に10校以上入る
日本再興戦略 2016 一第4次産業革命に向けてー		Ⅲ イノベーション・ベンチャー創出力の強化、チャレンジ精神にあふれる人材の創出等	
		2. 多面的アプローチによる人材の育成・確保等 2-3. 多様な働き手の参画 KPI：2017年末までに5,000人、2020年末までに10,000人の高度人材認定 iv) 外国人材の活用 ②外国人留学生の就職支援強化「国内での就職率を現状の3割から5割に向上」	1. イノベーション・ベンチャー創出力の強化 KPI：今後10年間で世界大学ランキングトップ100に10校以上
未来投資戦略 2017		Ⅱ Society 5.0 に向けた横割課題	
		3. 人材の育成・活用力の強化 KPI：2020年末までに10,000人、2022年末までに20,000人の高度外国人材の認定 iv) 外国人材の活用 ④外国人留学生の就職支援	
未来投資戦略 2018	Ⅱ. 経済構造革新への基盤づくり		
	2. AI時代に対応した人材育成と最適活用 2-3. 外国人材の活躍推進 KPI：2020年末までに10,000人、2022年末までに20,000人の高度外国人材の認定 KPI：2020年までに外国人留学生を30万人に（「留学生30万人計画」の実現） 新たに講ずべき具体的施策 ①外国人留学生等の国内就職促進のための政府横断的な取組 ②高度外国人材の受入、円滑化に向けた入国・在留管理制度等の改善		
成長戦略フォローアップ 2019	Ⅰ. Society 5.0 の実現		
	11. 外国人材の活躍推進 KPI：2020年末までに10,000人、2022年末までに20,000人の高度外国人材の認定 KPI：2020年までに外国人留学生倍増（「留学生30万人計画」の実現） 新たに講ずべき具体的施策 i) 高度外国人材の受入れ促進 ①外国人留学生等の国内就職促進のための政府横断的な取組 ②ビジネス日本語など教育プログラムの充実及び日本語教育の質の向上 ③高度外国人材の受入、円滑化に向けた入国・在留管理制度等の改善		

（出所）首相官邸「これまでの成長戦略について」に基づき筆者作成。

興プラン」[2. 雇用制度改革・人材力の強化]の⑦グローバル化に対応する人材力の強化には、「優秀な外国人留学生を2020年までに倍増（「留学生30万人計画」の実現）」という目標が示され、その方策として「優秀な外国人留学生獲得のための海外の重点地域を選定し、大学等の海外拠点の強化や支援の充実による戦略的な外国人留学生の確保を推進」（pp.37-38）と記されている。また、⑥「大学改革」に、スーパーグローバル大学を創設し、「世界大学ランキングトップ100に10校以上に入る」という目標が示され（p.36）、さらに「三. 国際展開戦略」の中堅・中小企業等への支援の項目に、留学生の日本就職を1万人とする目標が記載されている（pp.92-94）。

なお、同戦略の「優秀な外国人留学生獲得のための海外重点地域選定」という記述に呼応する形で、文科省に設置された「戦略的な留学生交流の推進に関する検討会」が「世界の成長を取り込むための外国人留学生の受入れ戦略」を策定し、受入れ重点地域として東南アジア、ロシア・CIS諸国、アフリカ、中東、南西アジア（特にインド）、東アジア（特にモンゴル）、南米、米国、中東欧を、重点分野として工学、医療、社会科学（法制度）、農学を選定している⁷。

「日本再興戦略改訂」（2014）においても、留学生30万人計画と大学等のグローバル化は、「大学改革/グローバル化等に対応する人材力の強化」の項目に並んで記載されていた。しかし2015年以降、大学等のグローバル化は、イノベーション関連の項目に移り、2017年以降は、「世界大学ランキングトップ100に10校以上」という目標が、成長戦略の本文中に見られなくなる。

他方、留学生の国内就職促進は、2014年以降、外国人材の確保／活躍支援の項に記されるようになり、日本再興戦略（2016）では、「国内での就職率を現状の3割から5割に向上」という具体的目標が示された（p.207）。それ以降の成長戦略では、留学生の国内就職に関する記述の増加が見られ、成長戦略フォローアップ（2019）では、ビジネス日本語や日本語教育の質の向上など具体的方策が記述され（pp.74-75）、30万人計画の中でも特に重視されていることがわかる。

成長戦略ポータルサイト「外国人材の活躍推

進」には、国際的な人材獲得競争が激化し、地域経済を支える人手不足が深刻化する中、国内企業のニーズに応じた外国人材が長期にわたり活躍できるよう、留学生の呼び込みから就職に至るまで一貫した対応を行うとともに、彼らが暮らしやすい社会を作っていくという政策目的が解説されている⁸。

（4）関係予算と施策・事業の変化

成長戦略における記述の変化を受け、30万人計画の関係予算と施策・事業はどう変化したのだろうか。表5は文科省の平成21年度と2019年度の予算書に記載された30万人計画関係予算を、30万人計画の5つの方策別に示している。

30万人計画の方策1の海外拠点と、2の日本留学試験の関連予算は、2009年度の6.9億円から2019年度には10.3億円に増加している。

方策3「大学等のグローバル化」関連では、2009年度は「グローバル30」に40.8億円が計上されていたが、2019年度には「スーパーグローバル大学創成支援事業（SGU）」に34億円、「大学の世界展開力強化事業」に13.1億円が配分されている。グローバル30は、英語による授業のみで学位が取得できるコースの設置、外国人教員の配置、留学生受入れのためのワンストップサービスなどを行う30の国際化拠点の整備事業として開始されたが、事業仕分けで予算が縮減し、初年度選定の13拠点のみの実施となった。

しかし2011年から、戦略的に重要な国・地域との間で、国際教育連携やネットワーク形成の取組みを支援する「大学の世界展開力強化事業」が開始され、さらに2014年には、日本再興戦略（2013）の「スーパーグローバル大学を創設」という記述を受けてSGUが開始され、37大学が選定され、初年度の予算は99億円であった。しかし、「世界大学ランキングトップ100に我が国の大学が10校以上」という記載が成長戦略から消えた2017年度には予算が63億円、2019年度には34億円と、大幅に減少している。

日本人学生の海外派遣には、2009年度に短期（協定派遣型）740人、長期（学位取得型）50人の予算が計上されていたが、2019年度は短期20,450人、長期371人と大きく増加した。また、

短期留学生受入れは1,800人から5,000人となり、留学生交流の予算は3.5倍に拡大している。

また、予算事項は異なるが、2019年度、「専修学校のグローバル化支援事業」に2億円が計上されている。

方策4の関連では、国費留学生の予算が220億円から185億円へと16%減少し、私費留学生の学習奨励費は、2016年度に渡日前の予約採用を重点化した「留学生受入れ促進プログラム」に変更されたが、予算は2009年度の78.9億円から2019年度は36.5億円へと半減している。2009年度に31.4億円が計上されていた私立学校の授業料減免学校法人援助、12.7億円が計上されていた留学生宿舍借上げ支援事業は、ともに廃止されている。

方策5の関連では、2009年度に「アジア人財資金支援構想事業」が文科省と経産省の共管で実施されていたが（表3）、事業仕分けのために2012年度に終了した。しかし、日本再興戦略（2016）の「国内就職率を3割から5割に」という目標達

成のため、2017年に「留学生就職促進プログラム」が開始され、2019年度に3.7億円が配分されている。

文科省の30万人計画の関連予算の総額は、2009年度の433.6億円から2019年度には392.8億円へと1割近く減額している。特に、奨学金や宿舍借上げなど、留学生に直接裨益する予算が35%の減額となっている。

文科省が2017年に主催した「高等教育機関における外国人留学生の受入推進に関する有識者会議」の報告⁹では、世界の熾烈な人材獲得競争に鑑み、「我が国として戦略的に招くべき学生」と「高度産業社会としての日本ないし日本文化に関心のある多様な学生」に分け、大学院での学位取得者が中心となる前者については、英語のみで学位取得が可能なコースの充実、教育の国際的質保証など教育プログラムの向上と奨学金供与が必要であり、後者については、授業料等コストの本人負担を求めることを提言してい

表5 文部科学省の留学生30万人計画の関係予算（2009年度，2019年度）

	2009 年度	2019 年度
事項	「留学生 30 万人計画」と大学の国際化	グローバル人材育成のための大学の国際化と学生の双方向交流の推進
方策 1	日本留学情報発信機能等の充実 348 百万円	日本留学海外拠点連携推進事業 450 百万円
方策 2	日本留学試験の拡充 342 百万円	日本留学試験他 583 百万円
	国際化拠点整備事業（グローバル 30） 4,081 百万円	スーパーグローバル大学創成支援事業 3,397 百万円
		大学の世界展開力強化事業 1,308 百万円
方策 3	留学生交流支援制度	日本人の海外留学促進事業（官民協働） 83 百万円
	・日本人学生の短期派遣 740 人	大学等の海外留学支援制度
	・日本人学生の長期派遣 50 人	・日本人短期（協定校） 20,450 人
	・外国人短期留学生受入れ 1,800 人 1,637 百万円	・日本人長期（学位取得型） 371 人
		・外国人短期留学生受入れ 5,000 人 8,017 百万円
方策 4	外国人留学生奨学金制度 33,107 百万円	外国人留学生奨学金制度 22,192 百万円
	・国費留学生 12,305 人 22,023 百万円	・国費留学生 11,276 人 18,545 百万円
	・私費学習奨励費 12,470 人 7,891 百万円	・私費（留学生受入れ促進プログラム） 7,800 人 3,647 百万円
	・授業料減免学校法人援助 3,138 百万円	
	留学生宿舍支援 1,269 百万円	
方策 5	留学生の就職支援の充実 38 百万円	留学生就職促進プログラム 370 百万円
事項		専修学校の人材養成機能の充実・強化
方策 3		専修学校グローバル化対応推進支援事業 196 百万円
合計	43,361 百万円	39,276 百万円

（出所）文部科学省「平成21年度予算主要事項」、「2019年度予算（案）主要事項」、「平成21年度我が国の留学制度の概要」に基づき、筆者作成。

る。文科省の予算配分においても、この提言に見られるように、奨学金は、将来の高度人材予備軍である優秀な留学生に絞って供与しようとする考え方が反映されていると見受けられる。

他省の動きとしては、外務省が日本再興戦略(2016)に示された優秀な留学生受入れのため、東南アジア・南西アジアの工学系トップレベル大学の人材を5年間で1,000人招へいし、留学やインターンシップ機会を提供する「イノベティブ・アジア」事業を2017年から開始している。また法務省は、第5次出入国管理政策懇談会の30万人計画実現に向けた提言に基づき、2009年の出入国管理法改正(2010年施行)において、日本語教育機関学生の「就学」の在留資格を「留学」に一本化した。また2016年には、日本語教育機関の急増に鑑み、日本語教育機関の告示基準を策定し、その後2回にわたり改定している。

(5) 30万人計画の政策的構図と評価の視点

以上の分析より、留学生30万人計画の目的は、計画自体に明示されていないものの、上位政策である成長戦略の記述から、最も重要な目的は、「優秀な留学生の増加」と「国内就職促進」を通じた高度人材の獲得であったと考えられる。

また、5つの方策のうち、大学等のグローバル化と優秀な留学生の増加との間には相互補完的關係が想定され、特に重視されていた。

成長戦略においても、当初は、30万人計画と大学等のグローバル化が、高等教育分野の主要なアジェンダとして推進されていたが、世界大学ランキングでの日本の大学の順位が上がらない中、大学等のグローバル化に関する記述は、2017年から成長戦略の本文中に見られなくなる。しかし、2019年度の文科省の予算には、SGUや大学の世界展開力強化事業が、30万人計画の予算と同じ事項に計上されており、相互補完関係に関する想定は、継続していると考えられる。

上記を踏まえ、次節では、30万人計画の主要な目的である「優秀な留学生の増加」と「留学生の国内就職」に関連する成果指標の分析から、その達成状況を検証するとともに、5つの方策、特に、相互補完関係が想定されていた大学等のグローバル化が、目的の達成にどのように寄与

したか、また、それ以外に目的の達成に影響を及ぼした要因は何であったかについて考察する。

3. 留学生30万人計画の主な成果指標の検証

(1) 優秀な留学生の増加

表6は、留学生30万人計画が開始された2008年と、30万人の目標が達成された2019年の学種／課程別の留学生数、入学経路、日本就職希望者及び就職者の割合を示している。

2008年から2019年にかけて、留学生と日本語教育機関で学ぶ就学生(2010年7月より留学生に一本化)の合計は、15.7万人から31.2万人へと倍増しているが、学種／課程別に見ると、専修学校(専門課程)在籍者は3.1倍、日本語教育機関在籍者は2.5倍と大きく増加しているのに対し、大学院で学ぶ者は1.6倍、学部で学ぶ者は1.5倍の増加にとどまる。2008年の計画発足当時、12万人だった留学生を30万人に増加するという計画では、「高度人材受入れとの連携」、「優秀な留学生の戦略的獲得」を謳い、高度人材の卵となる優秀な留学生を30万人に増やす(2.5倍にすること)を想定していたと考えられる。大学や大学院で学ぶ留学生がすべて優秀とは言えないものの、その増加が1.5~1.6倍にとどまることは、目標が十分に達成されていないことを示している。

表5に示す通り、予算削減から国費留学生が約千人減少した一方、私費留学生の割合は、2008年の89.8%から2019年には95.9%に上昇しており、留学生30万人計画は、私費留学生の増加によって達成されたと言っても過言ではない。国別では、2008年に中国出身者が58.8%と最も多く、次いで、韓国(15.2%)、台湾(4.1%)、ベトナム(2.3%)の順で、漢字圏(中国、韓国、台湾)出身者が78.1%を占めていたが、2019年には、中国(39.9%)、ベトナム(23.5%)、ネパール(8.4%)、韓国(5.9%)、台湾(3.1%)の順で、漢字圏出身者の割合が48.8%と縮小する一方、非漢字圏出身者が急増している¹⁰。

佐藤(2018)は、専修学校と日本語教育機関において、ベトナム、ネパールなどの非漢字圏

表6 学種／課程別の留学生数、入学経路、日本就職希望者及び就職者の割合（2008年、2019年）

	2008 年					2019 年					変化率
	留学生 数 (C)	海外から の直接入 学者割合	日本語教 育機関等 経由者 割合	日本就職 希望者 割合	日本就職 者割合	留学生 数 (D)	海外から の直接入 学者割合	日本語教 育機関等 経由者 割合	日本就職 希望者 割合	日本就職 者割合	
大学院 博士課程	32,666			49.4%	28.9%	53,089			56.3%	29.6%	1.63
大学院 修士課程		31.6%	57.4%	60.9%	31.4%		48.4%	33.5%	67.2%	33.7%	
学部	60,520	32.7%	45.8%	66.0%	34.9%	89,602	29.9%	49.1%	70.5%	42.0%	1.48
短期大学	2,117			52.7%	18.8%	2,844			82.5%	60.9%	1.34
高等専門学校	538				18.0%	506				4.0%	0.94
専修学校（専門課程）	25,753	28.1%	63.1%	61.3%	16.4%	78,844	12.1%	79.9%	75.9%	39.2%	3.06
準備教育課程	2,235				1.0%	3,518			46.3%	5.1%	1.57
高等教育機関の計 (A)	123,829			61.3%	25.3%	228,403			70.2%	36.9%	1.84
日本語教育機関 (B)	32,937					83,811			49.5%	6.9%	2.54
留学生総数 (A+B)	156,766					312,214			64.6%		1.99

(注1) 2008年には日本語教育機関で学ぶ者 (B) は就学生であり、留学生のカウントには含まれていなかったが、ここでは留学生数の変化とその内訳を示すために、留学生総数 (A+B) として示している。

(注2) 日本語教育機関等経由者割合は、日本語教育機関経由者割合と準備教育課程経由者割合の合計で、専修学校専門課程（日本語科を含む）の経由者は含まない。

(注3) 日本就職希望者割合は、進路に関する第1～第3希望の回答のうち、就職希望の回答の合計を回答者数で除して算出している。

(出所) 2008年と2019年の留学生数は、日本学生支援機構の平成20年度と2019（令和元）年度の「外国人留学生在籍状況調査結果」、2008年の日本語教育機関学生数（＝就学生数）は、「平成21年度我が国の留学制度の概要」に基づく。直接入学割合と日本語教育機関等経由者割合は、日本学生支援機構のデータを使った二子石（2021）の図1～3から算出。日本就職希望者割合は日本学生支援機構の平成19年度と平成29年度の「私費外国人留學生生活実態調査結果」（2年に1度実施）、卒業生中の日本就職者割合は日本学生支援機構の平成20年度と2019（令和元）年度の「外国人留學生進路状況・学位授与状況調査結果」に基づく。

出身者が増加していることを示し、その背景について、2011年の東日本大震災により、漢字圏出身者の帰国や来日中止が相次ぎ、危機感を抱いた日本語教育機関関係者が、これまで留学生が少なかった非漢字圏諸国での学生リクルートを強化したこと、その際、日本留学の利点として、資格外活動（アルバイト）の上限時間が週28時間（夏期、冬期、春期休暇中は1日8時間）と他国より長いことを強調し、これが、アジアの中で比較的貧しい国々の若者にアピールしたことを挙げている。さらに、「留学費用は借金しても日本でアルバイトすれば返済できる」といった甘言で勧誘する日本留学斡旋業者の増加や、日本における労働力不足が、「働きながら学ぶ」留学生増加の構造的要因になっていると分析し、長時間のアルバイトと非漢字圏という日本語習得上のハンディのため、日本語教育機関在籍中に十分な日本語能力を習得できず、大学進学をあきらめ、入学が比較的容易な専修学校へ進学する者が多いことを指摘している。

表6では、日本語教育機関等を経由して専修学校に入学する者の割合が、2008年の63.1%から2019年には79.9%に増加し、学部でも45.8%から49.1%に上昇している。この背景には、上述のような、日本語教育機関が非漢字圏からの「働きながら学ぶ留学生」のリクルートに果たす役割の大きさがある。

それでは、方策3の大学等のグローバル化は、優秀な留学生の誘致に、どのような役割を果たしたのだろうか。

英語による授業を実施している大学は、2008年度から2018年度にかけて、学部段階で26%から41.2%に、大学院（研究科）段階では29%から36.4%に増加している。また、同期間、英語による授業のみで学位が取得できる学科等がある大学は、学部段階では7大学8学部から42大学73学部（全大学の5.7%）に、大学院（研究科）では73大学139研究科から107大学269研究科（全大学の17.2%）へと拡大している¹¹。

表6に示す修士課程における海外からの直接入

学者の割合は、2008年の31.6%から、2019年には48.4%に上昇している。直接入学者が、英語で学位が取得できる学科の入学者とは限らないものの、英語による授業や英語で学位が取得できる学科の増設は、日本語習得という障壁をなくすことで、直接入学者の増加に寄与したと考えられる。

しかし、大学院で学ぶ留学生の増加が1.6倍にとどまったことは、中国を始めとする他の国々における大学のグローバル化や留学生受入れが予想以上のスピードで進んだことにより、当初、大学等のグローバル化施策・事業により意図された国際競争力の強化や、それを通じた優秀な留学生の誘致が、意図されたほどに達成できなかったことを示している。

(2) 留学生の国内就職

法務省の統計によれば、日本で就職した留学生は、2008年の11,040人から、2019年には30,947人へと2.8倍に増加している。日本就職者の最終学歴は、2008年には大卒が48.0%、大学院卒が30.8%、専修学校卒が16.5%であったが、2019年には大卒が41.4%、大学院卒が20.7%、専修学校卒が32.3%と、大卒、大学院卒者の割合が下がり、専修学校卒業者の割合が増加している。出身国別では、中国人留学生の割合が、2008年の69.3%から2019年には37.4%に下がり、代わってベトナム人（2019年就職者の22.7%）、ネパール人（同11.6%）など、非漢字圏出身者の割合が増加している¹²。

表6は、日本学生支援機構の「外国人留学生進路状況・学位授与状況調査結果」に基づく学種／課程別の日本就職者割合を示している¹³。2008年から2019年にかけて、日本就職者割合の伸びが大きいのが短期大学と専修学校で、2008年に18.8%、16.4%だったものが、2019年には60.9%、39.2%となっている。学部生の日本就職者割合も34.9%から42.0%へと増加傾向が見られ、高等教育機関の卒業者全体では、25.3%から36.9%となり、留学生30万人計画における日本就職促進策の効果が表れていると言えよう。他方、修士学生の日本就職割合は、31.4%から33.7%へ、博士学生では28.9%から29.6%へと、増加の幅が相対

的に小さい。「私費外国人留学生生活実態調査結果」に基づく日本就職希望者の割合は、2019年に修士課程在籍者で67%、博士課程在籍者で56%に上るところ、彼らが日本就職を希望しても、その実現を阻む要因が存在すると考えられる。

ディスコ（2020）が日本企業494社に行った調査では、外国人留学生に求める資質は、文系でも理系でも日本語力が第一位で、面接を日本語で行う企業も86%に上っている。英語で学位が取得できる学科で学ぶ留学生が大学院を中心に増加する中、このような企業側の日本語力の期待に応えられず、日本就職に失敗する留学生が少なくないことが要因の1つと考えられる。

文科省の「留学生就職促進プログラム」の取り組み状況に関する調査報告書（日本国際協力センター 2019：42-43）でも、企業は大学の想定より高い日本語レベルを期待しており、英語で学位が取れるコースが増加する中、大学入学時点での基礎的な日本語能力が不足し、ビジネス日本語まで到達できないケースがあることが、課題として挙げられている。

(3) 日本語教育機関の貢献と課題

上述のように、留学生の増加の中心は、専修学校と日本語教育機関であり、専修学校留学生の多くが日本語教育機関を経由して進学していることを考慮すると、日本語教育機関が30万人の数値目標の達成に果たした役割は大きい。

しかし、前節で見た30万人計画の方策や予算に、日本語教育機関に関するものがほとんど含まれておらず、法務省が日本語教育機関学生の「就学」の在留資格を「留学」に一本化したことと、日本語教育機関の告示基準を策定・改定したことがあるのみである。

このように日本語教育機関が、30万人計画の方策に十分に位置づけられていない背景には、専修学校について規定した学校教育法第124条に、「我が国に居住する外国人を専ら対象とするものを除く」という規定があり、外国人への教育を主とする日本語教育機関が、専修学校として認められてこなかったことが挙げられる。この結果、日本語教育振興協会（以下、日振協）が2014年に会員365校に行った調査では、学校法

人・準学校法人格を有する学校は30%にとどまり、株式会社・有限会社が57%に上り、7%が任意団体・個人・合資会社・特定非営利活動法人、6%が財団法人・社団法人・宗教法人・独立行政法人であった¹⁴。

日本語教育機関の設立は、当初法的規制がなく、行政上の指導や監督の対象ではなかったが、1998年の上海事件¹⁵を契機として、1989年に文部省（当時）、法務省、外務省3省共管の財団法人として日振協が設立され、国と共同で告示校（出入国管理法で留学生受入れを認められる日本語教育機関）の認定審査と定期的なモニタリングを行っていた（日振協 2010：21-23）。しかし、2010年の事業仕分けにより、この仕組みが廃止され、以後、法務省が文科省の協力を得て、告示校の認定審査を行っているが、認定後に、日本語教育の質や教育体制をモニタリングする仕組みが不十分と指摘されている（佐藤・堀江 2016：98）。

二子石（2019：64-65）は、告示校が2019年1月に665校と過去最大となったが、学校教育法の枠外である学校種別を温存したため、過剰な教育ビジネスが生じやすく、その結果、教育の質の低下をもたらすという構造が、1980年代から引き継がれていると指摘する。

日本語教育機関の7割が学校法人格を持たず、留学生30万人計画の中に積極的に位置づけられてこなかったという意味で、「留学生30万人計画が偽装留学生を増やした」という批判は的外れである。しかし、日本語教育機関が果たす役割の大きさにもかかわらず、30万人計画の方策の中に十分に位置づけてこなかった政策の空白が、「学びよりも働くことが中心の留学生」の増加につながった面は否定できない。

4. 結語

本稿では、留学生30万人計画について、文科省の関連文書や成長戦略における記述から、その主要な目的を「優秀な留学生の増加」と「留学生の国内就職」を通じた高度人材の獲得と特定し、これら2つの目的の達成状況の検証と、達

成に影響を与えた要因に関する考察を行った。

その結果、留学生30万人という目標は達成されたものの、増加の中心は専修学校や日本語教育機関で学ぶ留学生であり、将来高度人材となることが期待される大学や大学院留学生の増加は限定的で、優秀な留学生の増加が十分に達成されていないことが判明した。この背景には、留学生誘致や進学前の予備教育で重要な役割を果たす日本語教育機関が、留学生30万人計画の中に十分に位置づけられてこなかったこと、そのために、留学生の募集・選考が適切に行われないケースや、十分な日本語能力が身につかず、希望の進学ができないケースが少なくなかったことが挙げられる。これに鑑み、優秀な留学生の増加が実現できるよう、日本語教育機関による学生の募集・選考や教育に対して、国の支援とモニタリングを拡充していく必要がある。

また、30万人計画の主要な方策の1つとして位置づけられてきた大学等のグローバル化は、大学における英語による授業や英語による授業のみで学位が取得できる学科の拡大と、大学院を中心とする直接入学者の増加をもたらしたが、英語で教育を受けたがゆえに、企業の多くが求める日本語力に到達せず、日本での就職が困難となっている実態も明らかになった。このことから、今後は、大学院の英語コースで学ぶ留学生の日本就職支援策を強化する必要がある。特に修士課程は在籍期間が2年と短いことから、日本就職希望者には、来日前から日本語教育を行うなどの工夫が必要である。

留学生30万人計画の特徴の1つは、留学生に対する奨学金や支援経費を減らしながら、大学等のグローバル化などの環境整備を行ったことだが、そのことが、主にアルバイトで留学費用を賄う私費留学生の増加を招き、彼らが学業に集中できない状況を看過する結果につながっている。日本語教育機関在籍者は特に奨学金受給機会が少ないところ¹⁶、彼らのうち、勉学意欲が高い者に対する支援を拡充する必要がある。

最後に、本研究を通じた、政策評価への示唆を述べる。

留学生30万人計画の評価に当たって、目的が明記されていないことが課題となった。このた

め、文科省の「『留学生30万人計画の骨子』とりまとめ」の考え方を確認したところ、留学生交流の意義として14項目が列挙されていた。30万人計画は、文科省が原案を作り、その後、関係6省により公表されている。このプロセスで、6省間で計画の目的に関する合意形成が困難であった可能性がある。

他方、成長戦略における30万人計画関連の記述を追うと、成長戦略の各年のテーマに応じて、30万人計画関連の記述の位置、強調点や分量が変化していることが判明した。このことから、計画の目的を最初から確定せず、上位政策の方針に応じ、合致する目的や意義を示し、予算の獲得につなげるという行政側の戦略があったとも考えられる。

第2次・第3次・第4次安倍政権においては、経済の再生が最重要課題であり、多くの政策が、成長戦略に紐づけされていた。また、国内外の環境変化に応じ、成長戦略のテーマも変化した。このため、成長戦略の各年のテーマに応じて、強調点を変化させ、予算獲得を図るのは、行政の当然の対応であり、政策評価に当たっては、今後もこのような現象が起こる可能性を考慮し、政策の最も重要な目的を、上位政策との関係で、確認した上で評価を行う必要がある。

留学生30万人計画と大学等のグローバル化との間には、相互補完関係が想定されていたが、大学等のグローバル化により意図された国際競争力の強化や、それを通じた優秀な留学生の誘致は思うように進まなかった。その要因の1つに、近隣諸国などでの大学のグローバル化が、予想以上のスピードで進んだことが挙げられる。このことは、政策策定や評価において、外部条件を十分に考慮する必要性を示唆している。

大学等のグローバル化は、「優秀な留学生の誘致」につながると想定されていたが、入学時に日本語力を求めない英語コースの増加は、「留学生の国内就職」というもう1つの目的達成にブレーキをかける結果となった。関係施策／事業間の相互作用は、政策評価において注視すべき点である。

本評価での知見が、今後の政策評価の発展に

少しでも寄与することを期待する。

注記

- 1 例えば出井康博は、「留学生30万人計画」の歪んだ実態として、8割以上の日本語学校は“偽装留学生”頼みと指摘している（2019年4月23日付President Online, <https://president.jp/articles/-/28422>（2021年3月1日閲覧）。また、安藤海南男は、「只今激増中 エセ留学生と外国人犯罪」と題し、留学生30万人計画を満たすべくハードルを下げたばかりに、エセ留学生と外国人犯罪が増加したと主張する（マンスリーウイール 2017年6月号：223-233）。
- 2 首相官邸「第169回国会における福田内閣総理大臣施政方針演説」
<https://www.kantei.go.jp/jp/hukudaspeech/2008/01/18housin.html>（2021年3月1日閲覧）
- 3 首相官邸が2000年に発表した「IT基本戦略」では、2005年までに3万人のIT技術者を受け入れるという目標が示され、2003年の通商白書では、イノベーションを通じた経済成長のため、専門的技術的人材の受入れの必要性を提言している。
- 4 http://www.mext.go.jp/a_menu/koutou/ryugaku/_icsFiles/afieldfile/2019/09/18/1420758_001.pdf（2021年3月1日閲覧）
- 5 http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/houkoku/attach/1249711.htm（2021年3月1日閲覧）
- 6 https://www.mext.go.jp/a_menu/koutou/ryugaku/1284755.htm（2021年3月1日閲覧）
- 7 https://www.mext.go.jp/a_menu/koutou/ryugaku/1342726.htm
- 8 https://www.kantei.go.jp/jp/singi/keizaisaisei/portal/foreign_talent/（2021年3月1日閲覧）
- 9 https://www.mext.go.jp/component/a_menu/education/detail/_icsFiles/afieldfile/2017/08/21/1394116_002.pdf（2020年5月19日閲覧）
- 10 日本学生支援機構「平成20年度外国人留学生在籍状況調査結果」、「2019（令和元）年度外国人留学生在籍状況調査結果」に基づく。
- 11 文部科学省「大学における教育内容等の改革状況について（平成20年度）」、「平成30年度の大学における教育内容等の改革状況について」
- 12 法務省「平成20年における留学生等の日本企業等へ

- の就職状況について」、出入国在留管理庁「令和元年における留学生等の日本企業等への就職状況について」
- 13 日本学生支援機構の調査は、在籍校を通じた卒業時点の調査であるため、卒業後の在留資格変更を含む法務省／出入国在留管理庁の調査結果とやや異なる。
- 14 学校法人格を有する学校は、日本語科を有する専修学校であり、準学校法人格を有する学校は、都道府県に各種学校として認可された学校である。
- 15 日本語教育機関に入学金を支払ったにもかかわらず、就学ビザが発給されなかった中国の若者数百人が、上海の日本総領事館を取り囲んだ事件。
- 16 日本学生支援機構「平成29年度外国人留学生在籍状況調査結果」によれば、日本語教育機関在籍者の奨学金受給率は14.4%と、すべての学種の中で最も低い。
- 参考文献**
- 工藤和宏・上別府隆男・太田浩（2014）「日本の大学国際化と留学生政策の展開」、『日韓大学国際化と留学生政策の展開』、日本私立大学協会附置私学高等教育研究所、13-52
- 佐藤由利子（2010）『日本の留学生政策の評価－人材養成、友好促進、経済効果の視点から』、東信堂
- 佐藤由利子（2018）「移民・難民政策と留学生政策－留学生政策の多義性の利点と課題－」、『移民政策研究』、10：29-43
- 佐藤由利子・堀江学（2015）「日本の留学生教育の質保証とシステムの課題－ベトナム人留学生の特徴と送出し・受入れ要因の分析から－」、『留学生教育』、20：93-104
- 白石勝己（2019）「ポスト留学生30万人計画の行方～留学生30万人計画の総括と将来見通し」、『アジアの友』、540：2-15
- 総務省（2005）『留学生の受入れ推進施策に関する政策評価書』、総務省
- 総務省（2017）『グローバル人材育成の推進に関する政策評価書』、総務省
- 総務省（2019）『高度外国人材の受入れに関する政策評価書』、総務省
- 谷口吉弘（2011）『国費外国人留学生制度の成果・効果に関する調査研究』、文部科学省
- ディスコ（2020）「外国人留学生／高度外国人材の採用に関する調査」、ディスコ
- 西出順郎（2020）『政策はなぜ検証できないのか 政策評価制度の研究』、勁草書房
- 日本経済団体連合会（2004）『外国人受け入れ問題に関する提言』、日本経済団体連合会
- 日本国際協力センター（2020）『留学生の就職促進に関する周知及び調査研究（留学生就職促進プログラム成果報告書）』、文部科学省
- 日本語教育振興協会（2010）『日本語教育振興協会20年の歩み』、日本語教育振興協会
- 二子石優（2019）「『1990 年体制』成立を境にした日本国内の日本語学校の変移」、『一橋大学国際教育交流センター紀要』、1：55-68
- 二子石優（2021）「留学生30万人計画の達成とその実情を探る－留学生の入学経路と卒業後進路に関する一考察－」、『ウェブマガジン留学交流』、2021年3月号、Vol.120：42-60
- 山谷清志（2020）「評価体系、政策体系、行政体系－日本評価学会の20年の宿題－」、日本評価学会第21回全国大会基調講演資料
- 吉田文（2012）「2000年代の高等教育政策における産業界と行政のポリティックス：新自由主義・グローバル化・少子化」、『日本労働研究雑誌』、629：55-66
- 吉田文（2014）「『グローバル人材』の育成と日本の大学教育－議論のローカリズムをめぐる」、『教育学研究』、81（2）：28-39
- 「留学生30万人計画」関係省庁会議（2021）「『留学生30万人計画』骨子検証結果報告」、https://www.kantei.go.jp/jp/singi/kyouikusaiei/jikkoukaigi_wg/koutou_wg/dai8/siryoutu-3.pdf（2021年7月9日閲覧）
- （2021.7.28 受理）

**Outcomes and Problems of the Plan to Accept
300,000 International Students in Japan
– Analysis from the Relation with the Growth Strategy, Globalization of
Universities, and Japanese Language Education –**

Yuriko Sato

Tokyo Institute of Technology
yusato@tse.ens.titech.ac.jp

Abstract

The Plan to Accept 300,000 International Students achieved its target in 2019. In this paper, the Plan is evaluated considering its relation with the upper policy (Growth Strategy) and related measures/projects such as globalization of universities.

As the result of analysis, it was found out that globalization of universities led to the increase of international students in English-taught degree programs, especially at the graduate level. However, many of them found difficulty in finding employment in Japan because of the lack of Japanese language ability.

The main contributors to achieve the target of 300,000 international students are Japanese language schools. Since their student recruitment and educational activities have not been monitored enough by the government, it led to the increase of international students who are busy with part-time jobs and cannot concentrate in study.

Though the target number was achieved, the main purposes of the Plan (increase of excellent international students and their employment in Japan) are not fulfilled from this regard.

Keywords

International students, Globalization of universities, Internationalization of higher education, Employment of international students, Japanese language schools

【研究ノート】

実践的判断のプロセスとしての形成的アセスメント —J. デューイの価値評価論をめぐって—

西塚 孝平

東北大学大学院教育学研究科/日本学術振興会特別研究員

kohei.nishizuka.s1@dc.tohoku.ac.jp

要 約

近年の学校教育では、教授学習を改善するアセスメントの理論、すなわち形成的アセスメントの効果と有益性が叫ばれるようになってきた。しかし、理論形成に係る社会文化的視座が欠落しているがゆえに、理論的基盤の脆さと実践の適用可能性の狭さが批判されている。そこで本研究は、ジョン・デューイの価値評価論をめぐっての考察を通じて、社会文化的な形成的アセスメントの理論的枠組みを提案することを目的とする。その結果、特定のコミュニティの歴史を受け継ぐ教師が、情動（欲求、興味、努力）を働かせながら、観察を通じて事実判断と価値判断を行い、目論見を打ち立てることによって、実践的判断としてのフィードバックを生成させるという説明モデルが得られた。実践的判断のプロセスとして定式化される形成的アセスメントは、個人の変容と集団の変革という終わりなき目的・目標と教育活動の間を架橋し、教育的価値の相対化を通じて学校教育全体を創造していくポテンシャルを秘めた、コミュニティによる文化的実践といえる。

キーワード

形成的アセスメント、実践的判断、社会文化的アプローチ、ジョン・デューイ

1. はじめに

(1) 形成的アセスメント論の脆弱性

本稿の目的は、教授学習を改善するアセスメント、すなわち形成的アセスメント (formative assessment)¹ 概念を、ジョン・デューイの価値評価論 (theory of valuation) の立場から捉え直し、新たな理論的枠組みを提示することである。

形成的アセスメントとは、学習プロセス中に、生徒の現在の発達と目標のレベルを比較し、その間のギャップを埋める情報を、教育活動の質的な改善のために利用することである (Black &

Wiliam 1998, p.53)。教師や生徒が学習の進歩と課題を見極めた (アセスした) うえで、どこにいるのか (現在の発達水準の決定)、どこに向かうべきか (未来の発達水準の決定)、どうすれば良いのか (2つの水準の差を埋める助勢) への応答、つまりフィードバックを創出するとき、学習は前進する。このように観取した情報を、学習の価値を最終決定する総括的アセスメントの証拠として利用するのではなく、指導と学習の質をより良い状態へと変革させるために用いることの意義は、これまで国内外で数多く実証され、支持されてきた (Black & Wiliam 1998; Hattie

2009=2018; Nishizuka 2020)。

しかし、今から約30年前にサドラーが形成的アセスメント論を体系化させたものの(Sadler 1989)、今日まで複数の課題が残されたままとなっている。ここでは2点ほど指摘できるだろう。第1に、論理モデルがあまりにも単純すぎる。ブラックとウィリアムが言及しているように、人間による活動として実行されるアセスメントは、少なくとも言語や経験が基盤である以上、幅広い文脈による影響から逃れることはできない(Black & Wiliam 2018)。とりわけ私たちは、社会や文化²の中にある既存の思考・行動様式のパターン(文化的スクリプト)を学習している。エルウッドとマーフィーによると、この文化的スクリプトが、形成的アセスメントをローカルな理論として定義づける材料になるとともに、時として、アセスメントの目的や方法を歪め、葛藤を生む温床にもなる(Elwood & Murphy 2015)。こうした複雑な文脈を考慮せずにピア・アセスメント等のテクニックだけがまかり通ってしまうと、実践は字義通りに陥り、効果を得ることができなくなる(Black 2015; Marshall & Drummond 2006)。

第2に、教育目的・目標の矮小化である。前述の効果に関しては従来、テスト得点の向上が重要な証拠だとみなされ、政策決定にも反映されてきた(Hattie 2009=2018)。裏を返せば、形成的アセスメントは、数値測定ができる学習の範囲内での有効性が強調され、大学入試等のハイスティクスなテストに囚われた形であり続けている。ヒューズは、競争主義の渦に飲み込まれているアセスメントが豊かな学習の可能性を断ち切り、生徒の発達を抑制していることを非難している(Hughes 2014)。また、プレイヤーとクロスオードは、権力ある組織(主に政府機関)の言説や、テスト等の総括的アセスメントが正当な知識を決定づけると述べており(Pryor & Crossouard 2008)、教師自身がアセスメントの主導権を握ることができないでいる。こうして、説明責任を果たす普遍主義的理論を追求するあまり、かえってアセスメントの視野が狭まっていると同時に、さらにそのことを善とする力学が教室の中で働いているのである。

(2) 社会文化的アセスメント論の可能性

上記のような形成的アセスメントの理論的基盤の陥穽を、どのように克服できるだろうか。本稿では、アセスメントの社会文化的アプローチ(以下、社会文化的アセスメントと呼ぶ)に注目し、アセスメントを普遍的実践ではなく、コミュニティ³特有の文化的実践として再解釈するという切り口から、理論の再考に迫りたい。

社会文化的アセスメントはすでに、形成的アセスメント論が世界に普及し始める2000年代には提唱されていた(例えば、Elwood 2006; Gipps 2008; Shepard 2000)。彼らは、アセスメントを(社会、文化、歴史、言語、政治的な)文脈に依存した活動として再定義するために、知識詰め込みの個人的学習から問題解決中心の集団的学習への転換や、コミュニティの文化に注意を向けるように促している。特に、文化・歴史的活動理論(以下、活動理論と呼ぶ)は形成的アセスメント論を発展させる認識論だと言われている(Black & Wiliam 2018; Pryor & Crossouard 2008)。活動理論とは、人間と外界の対象がダイナミックに相互作用することでのみ、社会生活の組織化と発達は達成されたとする前提に立ち、彼らがどのような協働実践を重ねているのか、そして、直面する諸問題をいかにして自らの手で乗り越えていけるのかを明らかにする、「人と社会の弁証法的な発達過程を探究する理論」である(山住 2017, pp.i-ii)。実際、教師とアセスメントの関係性から発生する矛盾が、活動理論を用いて分析されてもいる(例えば、Burner 2019; Webb & Jones 2009)。

そもそも形成的アセスメントの核心には、活動理論に立つヴィゴツキーが提唱した最近接発達領域、すなわち現在の発達水準(生徒が1人でできること)と、未来の発達水準(他者の支援を受けて達成できること)の間にできる空間でフィードバックを駆動させるという発想がある(Allal & Ducrey 2000; Sadler 1989)。ところが、彼の理論は心理学や教育方法学の諸学問によって「個人主義的な系統学習論」へと還元された経緯がある(山住 2017, p.102)。アセスメント論にも外界を遮断する特徴が流れ込み、狭隘な理論と化してしまっていると考えられる。

しかし本稿では、最近接発達領域以外のヴィゴツキーの理論には立ち入らない。なぜなら、教育や学習に関する諸理論の概念化にあたっては、その認識論的な起源にまで遡って議論することが不可欠だからである (Engeström 2015=2020)。本稿では、形成的アセスメントの活動理論的解釈に挑戦する第一歩として、ヴィゴツキーより以前のジョン・デューイの教育論に着目する。ミエッティネンは、デューイと活動理論の関係について、媒介的道具を介して、心理の完全な内的存在（表象主義）と、心的プロセスが単なる外界からの反射活動にすぎないこと（反表象主義）を弁証法的に解消する企てと、それによる社会の再構築、つまりコミュニティの実践を変容させる志向性に関して、共通性がみられると主張する (Miettinen 2001)。グラスマンもまた、新たなコミュニティを創造する日常的な活動と、教育プロセスの中心としての最近接発達領域が一致する点に、デューイとヴィゴツキーの共鳴がみられることを示唆している (Glassman 2001)。したがって、個人と社会をつなぐ媒介物と、活動を鍵とした学習観からみて、デューイには活動理論の系譜が確実に流れているものと理解できる。また、最近接発達領域はアセスメントを活動理論的に解釈するにあたっての基礎であり、すなわち、デューイ理論を理解するための助けとして利用できる。

確かに、社会文化的アセスメントの先行研究は、デューイの精神機能と文脈の相互作用をしばしば引用してきた。例えば、教育目的・目標が本来持っている一時的かつ流動的な性質を指摘するとき (Sadler 1985)、教育を民主化するための教育哲学の始まりとしてアセスメントを位置づけるとき (Kucey & Parsons 2012)、最近接発達領域の横断において学習を統制可能とみる教育論を批判するとき (Egelandsdal & Riese 2020)、デューイの省察 (reflection) と活動中心の学習観は頼もしい理論的根拠を与えてきた。諸先行研究は教師が生徒の何をどのようにミ（見・看・観・診・視）るのが、指導内容や方法、生徒の学習スタイル、学習動機に作用することに特別な関心を払っている。また逆に、文脈的諸要因が発達の助勢に影響を及ぼすことを

理解したうえで、生徒と教師にとって複雑な思考の継起が、心的機能の次元ではなく二者間以上のコミュニケーションに基づく次元で理解すべきであることも議論されている。

このように形成的アセスメント論は、活動理論を基礎にして社会文化的な解釈を進めようとしている。だが、社会文化的アセスメントの論究は、いずれも形成的アセスメントと周辺因子の関連性のみが考究されており、概念がもつ性質を分析対象とはしていない。彼らが説く改善は、「主体である人間が道具を用いて対象に働きかける」という活動理論の一般モデルを用いて、「教師が形成的アセスメントを用いて学習支援を行う」のように定式化されている (Burner 2019; James et al. 2007; Webb & Jones 2009)。これは、形成的アセスメントが普遍的な手段として位置づけられ、概念自体が問い返されてはいないことを意味している。また、前述のデューイの理論に基づく論考では、経験や相互作用等の教授学習概念は持ち込まれているが、彼のアセスメントに関わる所説が形成的アセスメントの原理をどのように照らし出すのかまでは、直接検討されてきていない。

したがって、「個人主義的な系統学習論」としての最近接発達領域の考え方から抜け出すには、いちどデューイにまで立ち返り、デューイの理論の中で最近接発達領域を理解しなくてはならない。そしてこのとき、デューイが言明した価値評価 (valuation)、判断、観察等の関連概念を援用して形成的アセスメント概念を内部から対象化すれば、アセスメントの活動理論的解釈の推進に寄与できるのではないだろうか。

そこで本研究では、デューイの価値評価論をめぐる考察を通じて、形成的アセスメントの理論的枠組みを提案することを目的とする。第2章では、生徒の学習実態を把握することの意味と方法を、デューイが論じた観察と判断の視座から検討する。それを基に第3章では、実践的判断としてのフィードバックの原理を明らかにする。続けて第4章では、広範な文脈の中でフィードバックを機能させる戦略とその仕組みを、目論見とコミュニケーションの観点から整理する。終章では、デューイ論的解釈から得られる以上の

知見を総括し、社会文化的な形成的アセスメント論を構築していく。

2. 改善の基盤となる観察と判断の交互作用

(1) 観察：事実を特定するプロセス

デューイはアセスメントという言葉を直接使用してはいないが、「子どもを観る目の正確さ」(天野 1993、p.330)の通奏低音となる哲学を厚く展開させていた。というのは、何よりもデューイにとって、カリキュラム開発の始まりは教育目的・目標の設定ではなく、生徒が経験している教育プロセスの始点を特定し、生徒と学習対象の関係を正確に分析することにあつたからである(Dewey 1938=2004)。すなわち、教材や指導法の決定以前に、生徒の経験を知ることが教育の起点としていたのであった。

だとすれば、生徒が教師に改善の手がかりとなる情報を与えることでフィードバックが始まるとみたハッティの説明自体(Hattie 2009=2018)は正しいように思われる。「生徒の学習実態を把握する」とは、最近接発達領域における現在の発達水準を決するという意味である。教師が生徒の訴えや学習のポテンシャルを正確に把握できなければ、学習支援の内容と方法を根本的に見誤ってしまうだろう。それでは、「正確に把握する」とはどのような思考の手続きなのか。まずは、生徒の学習が順調に進んでいるのか、何につまずいているのかを教師が知ろうとすると、今まさに生徒はそのような状態であるという事実を特定しなくてはならない。

デューイは、この事実を構成する操作を観察と呼んだ(Dewey 1933=1955)。厳密にいうと、事実の構成には推理(inference)と推論(reasoning)から成る観念形成(ideation)と観念(idea)が必要になる(藤井 2010)。例えば今、ある高校生が小論文(対象の一部)の一節に「貧富の差が大きいから教育の質が下がる」と書いたとしよう。教師はこのことを視認し、なぜそれを問題として取り上げようとするのかを考え始める。1つは、視認したことと事実となりうる実態との関係を探る推理の操作があり、例えば「この説明は論

理的に誤った文章なのではないか」等と教師は考える。もう1つは、ここでは論証の概念が課題であると仮定すれば、そこから派生的に予想して、事実になりうる事柄を選択・特定しようとする推論の操作がある。この場合、「根拠に沿う主張を導出しているか」等の点が挙げられ、教師はその一節以外の記述にも目を配り始める。こうして、対象と観念形成を架橋する観察を通じて鍛錬される観念は、「貧富の格差と教育の質の間にある飛躍を埋めることができれば、論証を理解したといえるだろう」のように、予想や構想を含んだ、論理的解釈の道具となる。他方、観念が問題箇所を絞り込み、特定された事実は与件(data)と呼ばれる(Dewey 1933=1955, pp.140-141)。

したがって、観察のプロセスでは、問題を発見し、改善の対象とみる事実と、仮説的な結果と活動の方向性を予想する観念の2つが鍵を握る。しかしながら、ここにはさらなる疑問が残されている。そもそも何によって教師は当該部分に問題意識を抱けるようになるのか。いかにして推理や推論は出現するのか。その事実やその観念が最も相応しいことはどのように特定されるのか。観念はどのような形式で生徒に直接働きかけるのか。こうした疑問に答える思考の操作が次の判断だと考えられる。

(2) 判断：事実を意味づけるプロセス

判断とは、選択した事実の確実性、事実を正当化する諸原理・観念の適切性、そして、事実とそれに結びつく暗示(suggestion)との関係を、予想される結果との関連から決定する行動のことをいう(Dewey 1933=1955, p.123)。デューイは裁判官に倣い、判断の仕組みを3つの手順に分けて論じている(Dewey 1933=1955, pp.122-130)。

1つ目は、何よりもまず判断は疑惑と論争が火種となる。教師が小論文を目にして当該箇所に違和を覚える段階が、この部分に該当する。対象(事態)に怪しげな影を感じ取るとき、すなわち、異なる立場でも成立しうる解釈と出会い、問題の中心が意味するものを特定しようと準備する瞬間が、判断の発動条件になる。

2つ目は、妥当な諸事実と観念を選択し、その確実性と適切性を特定する。まず、諸事実の選

択についてデューイは、どのような事実が問題の中心となり、どれほど重要な意味をもつのかの識別は容易ではないとしている。

判断には、消去もしくは拒否が、解釈が、開示が、もしくは明るみに持ち出すことが行われねばならぬ。我々が最後の結論に到達するまでは拒否と選択とか試験的もしくは条件的でなければならぬ。我々は我々が意味を知る端緒であることを望みもし信じもするものを選択する。しかし、そのようなものが、そのものを是認し、そのものを包含する一事態を暗示しないならば、我々は与件を再構成する。(Dewey 1933=1955, p.126)

ゆえに事実とは、「視認できるあらゆる対象」ではなく、「観察プロセスを通じて諸特性や意味が付与された、選ばれた対象」だといえる。それでは、教師は小論文の中から事実となりうる事態を1つ1つ取り出し、観念にアクセスして、吟味にかけなくてはならないのだろうか。このことについてデューイは、「事態が含む重要な諸要素を把握する能力」としての呼吸 (knack)、機敏 (tact)、聡明 (cleverness)、洞察力 (insight)、鑑識眼 (discernment) が、「紛糾錯綜した事態がもつ種々の特性の関係を表示し意味する価値に関する勘 (sense) を働かせる」(Dewey 1933=1955, p.126) と述べている。判断者は経験的力量を即座に働かせ、問題解決に資する観念を通じて何らかの意味を付与できる対象に目星をつけていくというのである。本論では、このように事実となりうるものを与件化する操作を「事実特定」と呼ぶこととする。

もう1つの視角が、諸事実を最も的確に解釈できる観念の選択である。デューイは、観念の由来を暗示に求める。「価値に関する勘」と同様、暗示は「突如として心のなかにとびこむ」のであり、それは「或る種の既得の習慣的な理解様式」、「少くとも意味的な諸存在が引きだされ得る幾多の経験」、そのときの「一般文化状態」、「被暗示者自身の選択や願望や興味や或いは彼の直接の感情の状況」にも依存する (Dewey

1933=1955, pp.97-98, p.128)。つまり、何らかの合理的な契機が観念を出現させるのではない。観念の相応しさを決定するための外在的で絶対的な規準 (クライテリア、すなわち価値の程度を判断する物差し) は存在しないにもかかわらず、現実には相応しさの程度を特定せざるを得ないという堀田 (1965) の論考に基づけば、クライテリアの組成原理は人為的に構成され、構成した者にその責任が委ねられるものであると理解できる。このように、観念の相応しさを事実との対応関係から見極め、決定する操作を、本稿では「価値特定」と呼びたい。貧富の格差と教育の質への言及を見過ごしてしまう判断者は、観念の発見、およびクライテリアの生成と選択に失敗しており、諸事実を上手く整理し、結合できていない状態だといえる。

3つ目の手続きでは、観念を現実反映させる。価値特定によって裏付けを得た観念は、未来の発達水準を示唆する点で明らかに指示機能を内包しているため、教育活動では指導として現れるだろう。価値特定のプロセスは、将来の似た文脈でも活用できる諸法則や素材を明るみに出すことから (Dewey 1933=1955, p.129)、判断者もつクライテリアを強化させたり、逆にそれらを解体し再構成させたりもする。

以上より、観念を頼りにして事実を特定するプロセスには、事実を構成する観察と、それを意味づける判断との混成が認められる。さらにこのことから、次の点が指摘できる。すなわち、どの事実が生徒の学習において無視できないのかを、いかにそれが解決可能であるのかという視点から厳選する作業は、ともすれば特定される問題 (事実) を極めて局所的で非本質的なものにしかねないということである。教師が意識を向ける対象の範囲は小論文のみならず、生徒がこれまで教師に見せてきた正課と課外の学習全般、習慣的な言動や癖、性格、さらには家庭環境といった「教師がその生徒について知っている全て」にまで広がっている。このように対象に豊かさを求めることで、事実と観念には広さと奥行きが生まれ、全人的 (well-rounded) な観察と判断が可能になるのである。

とはいえ、前述までの観察と判断に関する整

理は、改善の観点からは直接論じられてはおらず、さらには、個人（主に教師）の認知的操作を取り扱うにとどまっている。そこで次章以降では、観察・判断と教授学習の改善との関連性を検討していく。具体的には、教師あるいは生徒が学習の良悪や可否といった価値を一体どのように決定し、指導と学習を最適に調整していくのかという、改善の働きかけの詳細なメカニズムを明らかにしていく。次章ではまず、価値の問題において重要なデューイの価値評価論の原理を整理したうえで、それを改善の働きかけのプロセスの中で再解釈する。

3. 実践的判断としての改善の働きかけ

(1) デューイの価値評価論

そもそも価値とは、デューイによれば現実から立ち上がってくるところの特性であり、超越的に絶対化されたものではない。それは、「状況的に可能と思われる、そして予見された完成体の一部になるものとしての特性」であり、「判断が所与の実存的な質に付与した影響力」（Dewey 1916, pp.515-516）である。また、磯野友彦翻訳の『評価の理論』に序文を寄せた帆足理一郎は、「人間行動の組織立のなかで、目ざされた行動の推進に役立つ」あらゆるものが価値であり、「自然と人間経験とのからみ合った局面または状況においてこそ、価値判断は発生する」（Dewey 1939=1957, p.2）と述べている。

よって価値は、観察を含む判断プロセスの中で発見され、人間による判断なくして価値はないとみる探究的立場に根ざしている。ゆえにデューイは、「価値」という名詞形ではなく「価値づける」という動詞形に着目したのである。

対象を価値づける方法について、デューイは次の2つに分類した。それは、対象の情動的・人格の性質に関わるものとして、代替できない本質的価値を主観的に「称賛する（praising, esteeming）」方法（valuing）と、対象を理智的・関係的性質の側面から捉え、比較・選択に使う道具的価値を客観的に「値踏みする（appraising, estimating）」方法（evaluation）である（Dewey

1939=1957, p.24）。前者の側面から対象を与件化する場合、判断者は感動や称賛といった受容感情で満たされる。例えば、上手く事が運んでいる生徒を教師が褒めたり、生徒が自己肯定したりするとき、価値の判断は実施されない。無論、別な文脈では称賛されない可能性もあるため、称賛の固有性は特定の文脈の中でのみ認められるものである。一方で後者の経験では、良否、善悪、優劣、可否、勝敗、美醜といった価値語を用いて比較し、選択しなくてはならず、目的・目標との距離、すなわち質へのこだわりが追求される。そのため、判断の初動期に感じる事態への疑念は、客観的値踏みですでに問題にしているといえる。客観的値踏みもまた、主観的称賛と同様に文脈が考慮される。例えば、生徒が100点満点の試験で30点であっても、良悪や可否はそのときの目的・目標との関係性から決定される。判断者が過去の諸経験や習慣に依存し、価値探究の視座を放棄してしまうと、客観的値踏みを誤ることになりかねない。

以上より価値評価では、本質的価値と道具的価値のいずれの場合であっても文脈の精密な分析が欠かせない。さらに価値評価は、同じ対象に対して2つの価値づける方法を同時に可能にする。例えば、同じ絵画を鑑賞したさい、神秘さから生まれる感動と、分析した後に結論づけられる感動の2つが得られる。実践に即して考えるならば、価値の探究方法を変化させることは、生徒の動機づけと学習の進歩、指導を調整することに関わるといえる。

(2) 実践的判断としてのフィードバック

とりわけ改善の働きに着目すれば、より重要なのは客観的値踏みである。デューイの論に依拠すると、客観的値踏みのプロセスは事実判断、価値判断、実践的判断の3つの段階に整理できるだろう。はじめに、観察可能な文脈が精密に分析され、形成しつつある（あるいはすでに形成させている）観念を参照して事実が特定される。その後、事実を文脈の中に定位させる「総合」と、その意味を増大させる「分析」が行われ、事実が客観的に何を語っているのかが意味づけられる（Dewey 1933=1955, p.133）。このように、

「直接には存在している状況」(Dewey 1939=1957, p.24) を明らかにする思考は事実判断と呼ばれる(森 1974)。先ほどの貧富の差の例を挙げると、「観察される事実からして、論理的思考ができていない」等となる。

また、観念形成では、「間接には作り出すことが意図され欲せられる未来の状況」(Dewey 1939=1957, p.24) が予兆されていることから、事実判断には、現在の状況と比較する手段、つまり道具的価値が付随してくる。デューイは、観念を参照して価値を事実が付与していく思考を価値判断と呼び、「われわれの願望、情感、喜びの形式を支配すべきことがらについての判断」(Dewey 1929=2018, p.265) だと定義した。先の例に続けば、「(観念に含まれる教育的価値と照合したときに、) 論理的思考ができていないのは良くない」等となる。

さらに、観念は指導の方向も指示しているため、価値は事実のさらなる推進力を生み出すものとなる。デューイは、「より良い価値の決定とは、それが何をするのがより良いかの決定に相当する(潜在力をもつ)」(Dewey 1916, pp.514-515) として、価値の本質が未来の行動やクライテリアによって与えられる点、価値判断には何がなされるべきかを決定する役割がある点にも言及している。それは、このように行動したほうが「より良い、より賢明である、より良識的だ、より正しい、より望ましい、より適切だ、より目的にかなっている」(Dewey 1916, p.505) 等の規範命題として出現する。そのような対象の不完全をより完全に近づける働きかけを、デューイは実践的判断(practical judgment)と呼んだ。この場合、「生徒には論理的思考を身につけるべきで、〔何らかの具体的な改善〕をするべきだ」等となる。これはフィードバックと呼ばれる活動にほかならない。

以上の論点は、次の2つに集約される。1つは、形成的アセスメントは実践的判断のプロセスとして説明できる。そしてこの考え方は、現在の発達水準を決定する事実判断、観念と事実の比較から発達の兆しを掴み取る価値判断、未来の発達水準を決定し、教授学習の進路を具体的に提供する実践的判断として、最近接発達領域に

対応している。もう1つは、事実の特徴に秘められた可能的成長に気づくことで初めて、フィードバックは生成される。ここでは、教育的価値をローカライズさせる重要性が指摘できる。デューイは信念との向き合い方について、「既知の善にしたがって行為する意志」ではなく「何が善であるかを知らうとする意志」に即して、行動の方向性を与える手段としての信念を仮説的に構築するべきであると論じている(Dewey 1929=2018, p.216)。斎藤もまた、既存の価値クライテリアの適用ではなく、判断者が所有する価値やクライテリアを疑うことにこそ価値判断の真価があると述べる(斎藤 1974)。つまり、実践的判断者は既存の価値と実践を照合するだけでなく、価値を創造する担い手となることで、活動を調整していくのである。

価値の受け手になることの弊害は、先行研究でも報告されている。特に知識創造学習において、判断の参照先(クライテリア)を厳格に規定すると、学習のニーズ把握とフィードバックの視野が狭まることが分かっている(James et al. 2007)。この課題に対処するために、教師と生徒でクライテリアを協働創造させるモデルが提案されている(西塚・有本 2019)。

ここまで、「生徒の学習実態を把握する」という現在の発達水準を中心とした考察を進めてきた。それでは、未来の発達水準という観点からみたとき、改善はどのように解釈できるのだろうか。もとよりデューイは教育目的・目標をどのように理解していたのか。また、いかにして観念を実践的判断へと具体化できるのだろうか。さらに、願望や情感等の主観的感情は実践的判断にどう作用するのだろうか。そこで最後に、目論見(end-in-view)と呼ばれるデューイの目的論に注目し、人間の意志がどのように改善行動に統合されていくのかを精査していく。

4. 目論見による改善の戦略設計

(1) 欲求の手段化

判断が人間の手によって行われるならば、その中に興味や欲求等の主観的な感情の影響が必

ず入り込んでくる。デューイは価値評価と欲求の関係を次のように述べている。

何となれば、何か欠けているものを生ぜしめることが必要である時にのみ、或いは、外的条件によって脅かされる何かのものを存在させておくことが必要である時にのみ、貴ぶ (prizing) とか欲する (caring) とかいう意味での評価 (valuation) が生ずるのであるから、即ち評価 (valuation) は欲求すること (desiring) を含むのである。(Dewey 1939=1957, p.30)

さらにデューイは、欲求を2つに大別している。それは、習慣や野生的な反応に身を委ねる衝動と、仮にそれに従うとしたらどうなるかの結果を見通したうえで批判的に再構成させた欲求であり、後者のみが考証たりうる知的な欲求であるとしている。このことをレオンチェフは、欲求は対象的であるときにのみ活動を統制できるという仕方で表現している。すなわち、欲求は「有機体の欠乏状態」(Leont'ev 1975=1980, p.73)ではなく、対象と出会い、現実の活動を方向づけ、調節する役割を果たす社会的状態を指している。

欲求以外にも、不足、危機、問題の解消には一定の忍耐力が求められる。生徒が抱えている困り感や、生徒の現実と理想の間にある障壁を乗り越えるとき、「容易ではないが、何とか解決したい」という対抗的な感情と葛藤することになる。デューイによれば、「欲求と反感のむすびついた特別の感情的状況」(Dewey 1913=1972, p.37)、あるいは「欲求に伴う緊張の本質」(Dewey 1939=1957, pp.32-33)が努力である。努力は、反射的な衝動から、抵抗に立ち向かう欲求へと発展させるような、活動を持続させるプロセスといえる。

加えて努力は、興味 (interest) を高めるエネルギーでもある。興味とは、「彼をして何か他の結果よりもある定まった結果を生ぜしめるように行動させるところの利害関係」(Dewey 1939=1957, p.34)である。何らかの利害をもつ人間は、現実(観察される対象や実態)と自らが望む理想(観

念)との間に生まれる食い違いを是正しようと試みる。天野の言葉を借りると、興味は「自我の側における事柄と自我によって働きかけられる事物や観念との同一化」(天野 1977, p.91)を欲求する働きかけであり、このときに努力は、興味の持続性を支えるのである。

したがって、外的に規定されてあるもの以外の教育目的・目標は欲求によって対象化され、厳選された観念といえるだろう。そして、この観念を改善のために利用しようと企てるとき、「欲望は、アイディア(観念)を刺激し、それに推進力を与え」、「アイディアは達成されるべき活動のなかにある計画や、またあるべき活動のための計画」へと成形されていく(Dewey 1938=2004, p.110)。デューイは、観念に基づいて現在の発達をより優れたものにするための手段とプロセスをデザインすることを目論見(end-in-view)ないしは観念力と呼んだ(Dewey 1939=1957, p.100)。つまり価値判断は、衝動から現実の活動を方向づける欲求へと至る経過の中で、可能な結果を予測するという教育目的・目標に根ざした思考を始めた瞬間に、実践的判断の契機を得るのである。

目論見が計画を指すとすれば、教育目的・目標は到達地点を単純に指示するものではなく、到達地点に向かうための方法となる。的(目論見)は投撃(学習)の可能性を高める手段であり、現在の行動のほうがむしろ目的だといえる。

厳密にいうと、目論見は現在の行動手段であり、現在の行動は遠い目的のための手段ではない。的が存在するから人が撃つのではなく、投げたり撃ったりすることがより効果的で重要かもしれないために的を設定するのである。(Dewey 1922, p.226)

例えば、「文章を論理的に書けるようになる」という目的・目標は、生徒の現時点の学習がより良く変化することが期待できるために打ち立てられるのであって、生徒の経験的な連続性を無下に学習の外側から課されるものではない。

以上から、目論見とは手段化された欲求であり、観念を実践的判断へと具体化していく手続

きである。また、生徒と教師が実践的判断を妥当かつ有益に働かせる仕組みともいえる。欲求によって対象化される観念は、目論見の構成を介して目的・目標を定めると同時に、推理や推論が、目的・目標の質を証明するクライテリアとなる。欲求、努力、興味、目論見の価値は、現実に関与させた結果をみて、それらが役立つものであったかを省察することで検証できる。

(2) コミュニケーションに埋め込まれたアセスメント

個人の主観的感情や観察、判断が目論見の構成要素であるとしても、それらは教師1人の思考で完結するわけではない。デューイは、コミュニティに帰属する人間に対して、文脈が無意識のうちに思考に作用するとしている。

何に価値があり、何に価値がないかについての、われわれの意識的な評価 (estimates) が、われわれが全くそれについて意識していない基準 (standards) に、どれほど多く依拠しているかについて、われわれは滅多に気づかない。しかし、一般に、われわれが探求や熟慮なしに、当然のことと見なしている事柄こそ、われわれの意識的思考を決定し、われわれの結論を決定する事柄であると言えるのである。そして、内省の水準の下にあるこれらの習慣こそが、他者との関係の絶え間ないやりとりのなかで形成された習慣なのである。(Dewey 1916=2000, p.34)

デューイは、コミュニケーションが「人びとが共同に事物を所持するにいたる仕方」であるとし、信念や願望、習慣、知識がコミュニティ内の共有物として創造、分有、継承されていくと考えた (Dewey 1916=2000, p.15)。このことを踏まえれば、「他者との関係の絶え間ないやりとり」、すなわちコミュニケーションが、専門知や身体知、経験、アイデンティティ、関係性のネットワークといった、教師 (集団) の生を支えている文化の維持と醸成に寄与していることは明らかである。教師は、帰属コミュニティが織

り成してきた歴史と、その中の様々な価値ある (valuing) 言動や価値クライテリアをスクリプトとして内面化していく。このように目論見の起源を辿ると、興味、欲求、勘、暗示、観念等は全てコミュニケーションによって創造されたものの、既存の文化から巡り巡って継承されてきたものとなる。そして、こうした目論見の背後にある「相互に連結した重層的なフィードバック・ループ」(Howe & Arimoto 2014, p.227) を加味する捉え方が、改善活動は「頭の中」の心的行為ではなく、「頭の外」の社会的活動として解釈されねばならないとする主張 (Elwood & Murphy 2015) を支持していることが分かる。

文脈的諸要因によって思考が制約を受ける以上、人間活動に見られる改善の原理と機構は、文化的価値システムの中でのみその全体像を掴み取ることができる。改善と文脈を統合的に理解する挑戦はすでに、いくつかの研究でなされている。有本と濱田によると、秋田県の小中学校では、全員の関わりと共同責任の感覚や教師のリーダーシップ等が、「どうすれば物事をよりよくすることができるかを問い続ける文化」を支えている。また彼らは、そこでのコミュニケーションに埋め込まれたアセスメントを、「地域が学校をシェイプする」という表現で指摘している (有本・濱田 2016, p.87)。つまり、学校を中心とした地域コミュニティの経時的かつ共時的なコミュニケーションによって、地域土着の言語やメタファー、歴史、伝統、規範、制度を水脈にして、実践的判断のプロセスが独自の形態で創発されているという。このように、フィードバック・ループはローカルなコミュニティでのみ創造・承認・伝承されるため、アセスメントを解釈する次元は個人から「集合的な生」へと引き上げなくてはならないのである。

5. 結論

本稿では、活動理論に立つデューイの議論にまで遡り、アセスメントを軸にして彼の理論を再解釈することで、形成的アセスメントの理論的枠組みを構成し直すことを試みた。つまりと

ころ、社会文化的な形成的アセスメントとは実践的判断のプロセスである。それは、個人の変容と集団の変革という目的・目標と教育活動の間を架橋し、教育的価値の相対化を通じて学校教育全体を創造していくポテンシャルを秘めた、コミュニティによる不断の文化的実践である。

具体的に、実践的判断のプロセスの要点は、次のように整理できる。教師はまず、事実となりうる対象（実態）を観察しようと興味をもつ。ここでは観察されうるものの範囲が決定される。次に、推論・推理による観念形成と、事実として取り出そうとしている対象との関係性を明らかにしたうえで、その対象の中から、生徒の学習の姿がどのようであり、何をしているのかという実態を選択的に与件化する。そのとき、事実特定と価値特定が、適切な事実と観念を決定する。その後、客観的にみて事実が何を意味するのか、生徒や教師は何かできて何ができていないのかを、分析と総合を通じて明らかにする事実判断を行う。この事実判断が改善の土台となるためには、明確なクライテリアを内包させている観念を持ち出し、現在と未来の発達水準の間にギャップを作る必要がある。その距離が学習の相対化・比較を可能にし、良悪や達成具合を客観的に値踏みする価値判断が実行される。この価値判断に作用する主観的感情は、一方では対象に注意を向ける興味となり、他方では価値判断に対して興味とともに衝動が働き、努力を介して「学習をどうしていきたいのか」の欲求を高める。目論見は、この欲求が推進力となり、観念に指導力を持たせる。この段階において、生徒や教師が実現できると思われる状態「…できる（はずだ）」が打ち立てられる。未来の発達水準が現在の行動をするために重要であることを正当化するために、「…するべきだ」という実践的判断が下される。実際に行われた実践的判断はその後、省察を経て、個人の経験、もしくはコミュニティの文化的価値システムを強化したり、再構成したりするように働きかける。

最後に、本稿の課題を2点述べる。第1に、本稿の議論と教育実践の分析の間にはまだ距離がある。例えば、教師は実践的判断の個々のプロセスを言語化するとも限らない。「生徒がつま

ず（学習が上手くできていない、それは良くないことだ）」等の言語的な特徴は、事実判断と価値判断の区別を曖昧にしている。また、教師間でもそうした言葉の意味を暗黙裡に了解し、コミュニケーションに使用してもいるだろう。そのため、教師と生徒の活動に注意を向け、実践的判断のプロセスを可視化する方法論を開発しなくてはならない。さらにいえば、そうした方法論以前に、デューイの教育論にも限界がある。デューイとヴィゴツキーの理論の相違性を見出したグラスマンによると、社会的な歴史とそこで生み出される道具、思考の資源となる経験と文化、学習の統制という面からみて、両者の見解は活動の使い方と教室で活動を使う理由が異なっている（Glassman 2001）。よって、近年の活動理論にまで至る系譜（Engeström 2015=2020）を丹念に追いかけて、本稿の知見に組み込み、議論を深化させる必要がある。

第2に、本稿では教育目的・目標の設定方法を論じたものの、その中身や、アセスメントとの関連性については触れていない。デューイにとって教育目的・目標は、「現存する状況を変えていく積極的過程」であり、「仕上げ、完成させ、洗練していく不断の過程」である（Dewey 1920=1995, p.133）。換言すれば、特定の学問領域の内容ではなく、経験を質的に変化させ、再構成させる継続的プロセスとしての成長が、真の目的・目標なのである。第1の課題と同様に、このことを実証面で解釈することが枢要となる。また、教育目的・目標を主題化すると、アセスメントの総括的側面が必然的に問題となってくる。本稿では、形成的アセスメント論とデューイの所論との関係性に関する言及にとどまっており、教育アセスメント論全体を取り上げたものではない。本稿の見解とデューイの所論によって、教室・測定学レベルの総括的アセスメントがどのように説明されうるのかを吟味していくこともまた、今後の課題である。

一口に形成的アセスメントといっても、人間による文化的実践として解釈する活動理論の立場と、従来のように、効果量（effect size）等を示して学習科学的に効果があると証明されてきた学習科学の立場があり、アセスメントの目的、

分析の射程範囲、駆動原理は明らかに異なる。また現時点では、この2つは理論と方法論をみても相容れない領域とされている (Engeström 2015=2020)。ゆえに、これまで蓄積されてきた学習科学側の考え方を活動理論の側が無条件に借用することは危険であり、アセスメントの社会文化的理論化には独自の認識論を用いるべきである。本稿の成果は、その認識論の1つとしてデューイ論的解釈を持ち出し、形成的アセスメントの理論的基盤を再構築したことにある。

付記

本研究はJSPS科研費20J20092の助成を受けたものです。

注記

- 1 本稿では、日本語の「形成的評価」という言葉を使用しない。「形成的評価」には、ブルーム (Bloom, B. S.) が考案した形成的テスト (特定の時期に行う小テスト)、小テストの合計を成績づけや評定に利用するという発想、指導や学習の後に実施する活動といった、心理測定学的な特徴を備えており、ペダゴジー、すなわち「教育理論、価値、エビデンス、正当性等に付随する言説であり、教えるという行為」 (Alexander 2008, p.47) の理論として理解されてきてはいない (Nishizuka 2020)。学習とアセスメントは本来、別々のプロセスではなく、ペダゴジーを介して複雑に絡まり合うものであるが (Baird et al. 2017)、国内ではその点が十分に議論されていない。「評価」に対する読み手のイメージを払拭するために、本稿ではアセスメントと表記する。ここでいうアセスメントとは、学習の状況と進歩に関する情報を収集し、整理するプロセスである (Black & Wiliam 1998)。
- 2 文化とは、活動を媒介する「コミュニティの共有物」とその体系のことであり、言葉、信念、規範、関係、役割等を指す (Dewey 1916=2000; Engeström 2015=2020)。
- 3 コミュニティとは、個々の成員が「共同の目的を知っており、それに関心をもっていて、それを考慮しながら各自の特殊な活動を調整する」集団であるこ

と、そして成員の間で、「共同に所有する事物」を創作し、所持する手段としてのコミュニケーションが成立している場である (Dewey 1916=2000, pp.15-16)。

参考文献

- 天野正輝 (1977) 「教育的興味概念の分析」、『東北大学教育学部研究年報』、25 : 83-108
- 天野正輝 (1993) 『教育評価史研究：教育実践における評価論の系譜』、東信堂
- 有本昌弘・濱田真 (2016) 「アセスメントによる評価の文化と秋田の小中学校：『インサイドザブラックボックス』の背後にある『グレーゾーン』にアプローチする」、『東北大学大学院教育学研究科研究年報』、65 (1) : 71-91
- 斎藤勉 (1974) 「教育と価値の関連構造の視点」、『日本デューイ学会紀要』、15 : 28-33
- 西塚孝平・有本昌弘 (2019) 「形成的アセスメント論におけるクライテリアの今日的意義：『深いESD』の実現に向けて」、『ホリスティック教育/ケア研究』、22 : 45-60
- 藤井千春 (2010) 『ジョン・デューイの経験主義哲学における思考論：知性的な思考の構造的解明』、早稲田大学出版部
- 堀田裕 (1965) 「デューイに於ける価値判断と事実判断との関連について」、『日本デューイ学会紀要』、6 : 16-22
- 森肇 (1974) 「デューイにおける価値判断と事実判断との関連について」、『日本デューイ学会紀要』、15 : 2-7
- 山住勝広 (2017) 『拡張する学校：協働学習の活動理論』、東京大学出版会
- Alexander, R. (2008) . *Essays on pedagogy*. London: Routledge.
- Allal, L., and Ducrey, G. P. (2000) . Assessment of—or in—the zone of proximal development. *Learning and Instruction*, 10 (2) , 137-152.
- Baird, J., Andrich, D., Hopfenbeck, T. N., and Stobart, G. (2017) . Assessment and learning: fields apart? *Assessment in Education: Principles, Policy & Practice*, 24 (3) , 317-350.
- Black, P. (2015) . Formative assessment: an optimistic but

- incomplete vision. *Assessment in Education: Principles, Policy & Practice*, 22 (1), 161–177.
- Black, P., and Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5 (1), 7–74.
- Black, P., and Wiliam, D. (2018). Classroom Assessment and Pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25 (6), 551–575.
- Burner, T. (2019). Aligning assessment and learning using cultural-historical activity theory. *Assessment Matters*, 13, 91–112.
- Dewey, J. (1913). *Interest and Effort in Education*. NY: Houghton Mifflin Company. (1972、杉浦宏 (訳) 『教育における興味と努力』、明治図書)
- Dewey, J. (1916). The Logic of Judgments of Practise. *The Journal of Philosophy, Psychology and Scientific Methods*, 12 (19), 505–523.
- Dewey, J. (1916). *Democracy and Education: An Introduction to the Philosophy of Education*. NY: Free Press. (2000、川村望 (訳) 『民主主義と教育』、人間の科学社)
- Dewey, J. (1920). *Reconstruction in Philosophy*. NY: Holt and Company. (1995、川村望 (訳) 『哲学の再構成』、人間の科学社)
- Dewey, J. (1922). *Human Nature and Conduct: An Introduction to Social Psychology*. NY: The Modern Library.
- Dewey, J. (1929). *Quest for Certainty: A Study of Relation of Knowledge and Action*. London: George Allen & Unwin Ltd. (2018、加賀裕郎 (訳) 『確実性の探求』、東京大学出版会)
- Dewey, J. (1933). *How we think: a restatement of the relation of reflective thinking to the educative process*. Boston: Houghton Mifflin. (1955、植田清次 (訳) 『思考の方法：いかにわれわれは思考するか〔普及版〕』、春秋社)
- Dewey, J. (1938). *Experience and Education*. NY: The Macmillan Company. (2004、市村尚久 (訳) 『経験と教育』、講談社学術文庫)
- Dewey, J. (1939). *Theory of Valuation*. Chicago: University of Chicago Press. (1957、磯野友彦 (訳) 『評価の理論』、関書院)
- Egelandstad, K., and Riese, H. (2020). Never mind the gap: Formative assessment confronted with Dewey's and Gadamer's concept of experience. *European Journal of Education*, 55, 91–104.
- Elwood, J. (2006). Formative assessment: possibilities, boundaries and limitations. *Assessment in Education: Principles, Policy & Practice*, 13 (2), 215–232.
- Elwood, J., and Murphy, P. (2015). Assessment systems as cultural scripts: a sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, 22 (2), 182–192.
- Engeström, Y. (2015). *Learning by Expanding: An Activity-theoretical Approach to Developmental Research (2nd Edition)*. NY: Cambridge University Press. (2020、山住勝広 (訳) 『拡張による学習：発達研究への活動理論からのアプローチ〔完訳増補版〕』、新曜社)
- Gipps, C. (2008). Socio-cultural aspects of assessment. In H. Wynne (Ed.), *Student assessment and testing: Vol. 1*, 252–291. CA: Sage.
- Glassman, M. (2001). Dewey and Vygotsky: Society, Experience, and Inquiry in Educational Practice. *Educational Researcher*, 30 (4), 3–14.
- Hattie, J. (2009). *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge. (2018、山森光陽 (監訳) 『教育の効果：メタ分析による学力に影響を与える要因の効果の可視化』、図書文化)
- Howe, E. R., and Arimoto, M. (2014). Narrative Teacher Education Pedagogies from Across the Pacific. In C. Craig & L. Orland-Barak (Eds.), *International Teacher Education: Promising Pedagogies Advances in Research on Teaching Part A*, 217–236. NY: Emerald.
- Hughes, G. (2014). *Ipsative Assessment: Motivation through Marking Progress*. UK: Palgrave Macmillan.
- James, M., McCormick, R., Black, P., Carmichael, P., Drummond, M. J., Fox, A., ... and Wiliam, D. (2007). *Improving learning how to learn: Classrooms, schools and networks*. London: Routledge.
- Kucey, S., and Parsons, J. (2012). Linking Past and Present: John Dewey and Assessment for Learning. *Journal of Teaching and Learning*, 8 (1), 107–116.
- Leont'ev, A.N. (1975). *Activity, Consciousness,*

- Personality*. Moscow: Politizdat. (1980、西村学・黒田直実 (訳)『活動と意識と人格』、明治図書)
- Marshall, B., and Drummond, M. J. (2006) . How teachers engage with Assessment for Learning: lessons from the classroom. *Research Papers in Education*, 21 (2) , 133–149.
- Miettinen, R. (2001) . Artifact Mediation in Dewey and in Cultural-Historical Activity Theory. *Mind, Culture, and Activity*, 8 (4) , 297–308.
- Nishizuka, K. (2020) . A Critical Review of Formative Assessment Research and Practice in Japan. *International Journal of Curriculum Development and Practice*, 22 (1) , 15–47.
- Pryor, J., and Crossouard, B. (2008) . A sociocultural theorisation of formative assessment. *Oxford Review of Education*, 34 (1) , 1–20.
- Sadler, R. D. (1985) . The Origins and Functions of Evaluative Criteria. *Educational Theory*, 35 (3) , 285–297.
- Sadler, R. D. (1989) . Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Shepard, L. A. (2000) . The Role of a Learning Culture. *Educational Researcher*, 29 (7) , 4–14.
- Webb, M., and Jones, J. (2009) . Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 16 (2) , 165–184.

(2021.7.6 受理)

Formative Assessment as a Process of Practical Judgment – Return to Theory of Valuation by John Dewey –

Kohei Nishizuka

Tohoku University, JSPS Research Fellowship for Young Scientist
kohei.nishizuka.sl@dc.tohoku.ac.jp

Abstract

In recent years, educators have been crying out for an effective and practical theory of assessment to improve teaching and learning, i.e., formative assessment. Such assessment has been criticized, however, for its fragile theoretical foundation and limited applicability to practice, both of which arise from the failure to incorporate socio-cultural perspectives when constructing theory. This study, therefore, proposes a new theoretical framework for socio-cultural formative assessment, based on a discussion of John Dewey's theory of valuation. An explanatory model was thus obtained in which feedback as a practical judgment is generated by teachers living in a particular community. These teachers make factual and value judgments informed by their observations and create ends-in-view while exercising emotions (desire, interest, and effort). Formative assessment, which can be treated as a process of practical judgment, is a cultural practice of a specific community with the potential to qualitatively change the entire school education system. This can be achieved by relativizing educational values, bridging gaps between educational activities and ends-in-view for the transformation of both the individual and the group.

Keywords

formative assessment, practical judgment, socio-cultural approach, John Dewey

日本評価学会 第22回全国大会のご案内

企画委員長

齊藤 貴浩（大阪大学）

第22回全国大会実行委員長

田中 啓（静岡文化芸術大学）

日本評価学会第22回全国大会については、昨今の新型コロナウイルス感染症（COVID-19）の影響等に鑑み、オンラインにて開催いたします。詳細は学会メーリングリストおよびホームページにてご案内いたします。

記

1. 日程：2021年12月4日（土）、12月5日（日）
2. 方法：オンラインにて開催（前回と同様にZoomを用います）

以上

～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。投稿の締め切りは9月末日（翌年3月刊行）及び3月末日（9月刊行）です。

ご興味をお持ちの方は投稿規定・執筆要領・査読要領、並びに原稿見本をご参照のうえ、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 編集委員会事務局

E-mail： journal@evaluationjp.org

『日本評価研究』第21巻第2号

2021年 9 月30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
〒113-0034 東京都文京区湯島3-31-1
中川ビル5階
一般財団法人行政管理研究センター内

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. 21, No. 2, September 2021

CONTENTS

Preface

Towards active management of JES that can provide a larger number of professionals
interested in evaluation research with a more stable and interactive community

Iwao Oshima

Special Issue: Issues on Evaluating the SDGs

Editor's Note: Issues on Evaluating the SDGs

Kaoru Hayashi

Evaluation Capacities to Advance Sustainable Development for All

Indran A. Naidoo

Evaluation Capacity Development in SDGs Era

Keiichi Muraoka, Ruiko Hino

A Consideration on Evaluation for Conceptual Targets of SDG4:
Shift from Measurability to Evaluability

Aki Yonehara

Evaluation of the "Goal 7" of the SDGs

Kaoru Hayashi

Interpretation and application of DAC New Evaluation Criteria

Masayuki Eguchi

Evaluation Use in Framing New SDG Projects:

Case Report of an African Youth Entrepreneur Capacity-building Project

Masafumi Nagao

Research Note

Social Impact Measurement as a Type of Program Evaluation

- Challenges and Opportunities - Ken Ito, Masatoshi Tamamura, Junta Ueno

Outcomes and Problems of the Plan to Accept 300,000 International Students in Japan

- Analysis from the Relation with the Growth Strategy, Globalization of Universities,
and Japanese Language Education - Yuriko Sato

Formative Assessment as a Process of Practical Judgment

- Return to Theory of Valuation by John Dewey - Kohei Nishizuka

Announcement for the 22nd Annual Conference