

4 研究開発評価の多様性

三菱総合研究所

要 旨

研究開発評価において、単に研究開発成果の達成のみを問うのではなく、多面的な観点で評価が行われている事例が見られる。

研究開発はそれぞれ多様な目的や期待される成果をもっており、その性格に応じて評価方法も自ら異なってくる。また、研究開発評価は一面的・画一的に実施されると、それに対して過剰な適応が行われるために、かえって研究開発の目的の達成が阻害される場合がある。例えば、短期的な成果を過度に求めるために、人材育成など中長期的に研究基盤を維持するための活動が損なわれることや、失敗のリスクが少ない小粒なテーマに偏重したりすることが挙げられる。

各国の事例においても、こうした点に配慮した特徴ある評価基準が設けられているが、望ましい手法については試行錯誤が行われている。

I 多面的評価の背景

研究開発には分野や基礎、応用といった性格の違いがあるが、国による研究開発ではさらにその目的が、学術振興、政策課題の解決、科学技術の基盤整備と多様なものになっている。さらに、研究開発課題の設定についても、研究者の自由な提案によるものから、領域やテーマが予め特定されたものもある。

こうした多様な目的や期待される成果があり、それを実現するために評価が行われるため、対象・時期・観点が異なる様々な評価方法がとられている。例えば、学術振興を目的とした研究開発では、学術的卓越性が高いと思われる課題を研究者によるピア・レビューによって選定し、実施は各研究主体に委ねられる。一方、特定の政策課題の解決を目的とした研究開発では、その目標の実現可能性に重点をおいた審査と評価が行われる。

研究開発評価が一面的・画一的に実施されると、それに対して過剰な適応が行われるために、かえって研究開発の目的の達成が阻害される場合がある。

たとえば、科学技術・学術審議会研究計画・評価分科会研究評価部会では、『研究開発評価システム改革の方向性について（審議のまとめ）』（平成21年8月4日）の中で、評価の観点・基準・視点について、表1に示す5つの視点を示している。

表1 評価システムの当面講ずべき改革の方向性～評価の観点・基準・視点

(1) 研究開発に適した評価の観点	現行の「必要性」、「有効性」、「効率性」に代わる研究開発に適したよりふさわしい評価の観点があるのではないか。
(2) 研究開発の性格に応じた多様な評価基準	研究開発の性格に応じて、評価方法は異なるはずであり、性格に応じた評価基準を明確化する必要があるのではないか。
(3) 研究活動を支える組織、次世代の人材を育成する組織やプログラムの役割を重視する評価の視点	第3期科学技術基本計画では、「モノから人へ、機関における個人の重視」の基本姿勢を示し、人材育成や競争的な研究環境の醸成が進められたが、一方で、現場では、個人を支える基盤の脆弱や次世代の若手研究者が育つ環境が整備されていない状況が見受けられることから、個人を支える機関の役割を重視するような評価の視点、あるいは次世代の人材育成を重視する評価の視点が必要ではないか。

(4) 研究コミュニティの活性化を促進する評価の視点	独創的な研究や新たな研究領域の開拓を促す価値観の融合を活性化するため、既存の研究グループや研究プロジェクト等の枠組みを超えた、研究コミュニティの動的な研究展開を促進しうる評価とするべきではないか。
(5) 世界的な視点での評価	世界的な視点での評価について、世界的なベンチマークの活用等、研究開発の特性に応じた世界水準の評価方法など、我が国にふさわしい評価方法を明確化する必要があるのではないか。

(出典) 科学技術・学術審議会研究計画・評価分科会研究評価部会『研究開発評価システム改革の方向性について（審議のまとめ）』（平成21年8月4日）を基に筆者作成

これらで指摘されているのは、研究開発評価には次の2つの「多様性」が必要であるという点である。すなわち、研究開発評価の多様性とは、個々の研究開発で評価の観点を増やすことにとどまるものではない。

- 目的、対象、性格、内容、規模等が研究開発によって多様であるため、研究開発評価も画一的ではなく、それぞれに応じた観点で実施する必要がある。
- 研究開発そのものを評価するだけでなく、そうした研究開発が継続的に実施できるために人材等の基盤も含めた多様な観点も評価する必要がある。

Ⅱ 研究開発評価の多様性の事例⁽¹⁾

ファンディング機関による研究開発課題の審査（事前評価）でどのような多面的評価の取組があるかを整理する。ただ、研究開発資金の配分方法は国や組織によって異なっており、ファンディング機関による研究開発課題の審査（事前評価）の位置づけ、役割も異なっていることには留意が必要である。

1 日本

我が国では、大学や独立行政法人に対しては運営費交付金として機関に基盤的な研究資金が配分された上で、研究開発課題に対しては競争的資金等による配分が行われている。

我が国の「国の研究開発評価に関する大綱的指針」では、評価の観点としては、以下のよう
に必要性、効率性、有効性の3点を挙げており、これが各府省における研究開発評価でも基本
となっている。

評価は、「行政機関が行う政策の評価に関する法律」において示されている政策評価の
観点も踏まえて、必要性、効率性、有効性の観点から、また、対象となる研究開発の国際
的な水準の向上の観点から実施する。また、これらの観点の下、研究開発の特性や評価の
目的等に応じて、適切な評価項目及び評価基準を設定して実施する。

(1) 日本学術振興会による科学研究費補助金

科学研究費補助金は研究者の自由な発想と研究意欲を源泉として行われる知的創造活動である
学術研究を支援する制度である。学術研究は当初の目的を達成できないこともあり、研究の

(1) 文部科学省『評価者の視点を通じた研究コミュニティの活性化に関する調査・分析』2010. を参考に作成。

経過そのものや時には失敗さえ価値を持つことがあるという特徴を持っている⁽²⁾ため、特定の応用の実現を明確な目標として行う、いわゆる出口指向の研究とは異なる評価の方法が必要と考えられる。

科学研究費補助金は、予めテーマを定めるのではなく、公募による研究者からの研究計画の申請を審査している。審査は、研究者が審査員となって書面で審査し、その結果を参考に審査会の合議で決める2段階審査方式をとっている。この審査員は、日本学術振興会の学術システム研究センターに所属するプログラムオフィサー（Program Officer）が候補者を選定し、審査員選考会を開いて決定する。なお、選定された審査員は通常2年務めるが、各審査員の審査結果は事後的にチェックされており、1年目の審査内容に問題が認められた場合には、2年目には審査依頼を行わないことがある。

主要な種目である基盤研究（A・B・C）の第1段階審査の基準は次のように示されている。すなわち、(1)では社会的な課題の解決ではなく学術的な重要性が問われており、(4)で波及効果として学術的な波及効果とともに社会に与えるインパクト・貢献が考慮されている。また、(5)については、体制や申請時までの実績から研究遂行能力を評価しているが、当該研究開発課題の成果の確実性を直接問うていない⁽³⁾。

- (1) 研究課題の学術的重要性・妥当性
- (2) 研究計画・方法の妥当性
- (3) 研究課題の独創性及び革新性
- (4) 研究課題の波及効果及び普遍性
- (5) 研究遂行能力及び研究環境の適切性
- (6) 研究計画と研究進捗評価を受けた研究課題の関連性

(2) 科学技術振興機構の戦略的創造研究推進事業

科学技術振興機構が実施する研究開発は課題解決型の基礎研究であり、学術研究とは異なり、ミッションがあることが特徴である。

戦略的創造研究推進事業として、個人対象・公募型の「さきがけ」、グループ研究・公募型の「CREST」、グループ研究・総括実施型の「ERATO」等、実施形態が異なる研究を行っている。いずれも、文部科学省が示した戦略目標を受けてJSTが研究領域と研究総括を定めて研究を実施する。

「さきがけ」では、選考にあたっての基準として以下が挙げられている⁽⁴⁾。

- a. 戦略目標の達成に貢献するものであること。
- b. 研究領域の趣旨に合致したものであること。
- c. 提案者自身の着想であること。
- d. 独創性を有していること。
- e. 研究構想の実現に必要な手掛かりが得られていること。
- f. 今後の科学技術に大きなインパクト（新技術の創出、重要問題の解決など）を与える可能性

(2) 科学技術・学術審議会学術分科会『学術研究における評価の在り方について（報告）』（平成14年2月14日）

(3) 独立行政法人日本学術振興会「平成23年度科学研究費補助金第1段階審査（書面審査）の手引」2010.

(4) 独立行政法人科学技術振興機構「平成23年度戦略的創造研究推進事業（CREST、さきがけ）研究提案募集のご案内〔第2期募集要項〕」2011, p.46.

を有していること。

- g. 研究が適切な実施規模であること。

さらに、さきがけには「大挑戦型」が設けられている。これは、成功した場合には飛躍的、画期的な成果が期待できる研究（いわゆる、大化けする研究）であって、実現の可能性の観点からは明確な見通しが得難いハイリスク研究を積極的に採択するものである。さきがけ通常型の選考プロセスに加え、幅広い視点から、提案の可能性・期待性の審査を行うとしており、基準に以下が含まれている。また、中間、事後評価では、ハイリスク研究に挑戦したことを前提とした評価を行うとしている⁽⁵⁾。

- h. 実現の可能性の観点からは明確な見通しが得難いが、成功した場合に飛躍的、画期的な成果が期待できること。

2 米国

米国の研究費の配分は、我が国の国立大学や独立行政法人に対する運営費交付金に相当する仕組みではなく、競争的な方法が主である。

(1) NSF (National Science Foundation)

NSFは、医学を除くあらゆる基礎科学・工学を支援している。他の連邦機関と異なり、研究者雇用や、研究所の運営には直接関与していない。活動の大部分は有期のグラント（補助金）交付によるものである。NSFの支援額は、米国の大学において行われる連邦政府の支援による基礎研究支出の約20パーセントにあたる⁽⁶⁾。NSFの組織は学術分野別に7つのDirectorateに分かれており、さらにその下に5つ程度のDivisionに分かれている。

プロジェクト選定の評価においては、主に次のような手順を踏む⁽⁷⁾。

1. NSF Program Officerが予備審査（要件、書式等の審査）。Peer reviewer最低3名を選定。
2. Peer Reviewer（external reviewer）が提案書を評価・分析し、NSF Program Officerに提示。
3. NSF Program Officerが外部評価者の評価を総括・審査。さらに、当該評価以外の要素を検討（例：分野における状況変化、予算状況、他の案件とのバランス等）。その上で、採択・不採択を提示。
4. 通常、Divisionレベルで最終決定。Program Officerによる評価に、Division Directorが同意して決定。

Program Officerは、常勤であり、いずれも当該研究分野において高い専門性を持つ者であるが、長期継続的雇用形態の者と、rotatorと呼ばれる任期付き雇用の者がいる。後者は大学や政府・非営利研究機関等から派遣されている科学者、エンジニア、または教育者がいる。

(5) 独立行政法人科学技術振興機構「平成23年度戦略的創造研究推進事業（CREST、さきがけ）研究提案募集のご案内〔第2期募集要項〕」2011, p.46.

(6) US NSF - About - NSF at a Glance <<http://www.nsf.gov/about/glance.jsp>> [last accessed: 2012/2/3]

(7) US NSF - Merit Review: Phase II: Proposal Review and Processing <<http://www.nsf.gov/bfa/dias/policy/meritreview/phase2.jsp#select>> [last accessed: 2012/2/3]

Peer Reviewerの選定に際しては、評価対象分野に関する専門・関連知識が挙げられているが、さらに、可能な限り同一評価グループ内に多様な（所属、年齢層等が異なる）人員を含めることとしている。

NSFの評価基準は1998年以来、メリット・レビューとして知られており、表2で示す2つが挙げられている。プログラムによってはさらに評価基準が追加される場合もある。

表2 NSFにおける提案の評価基準

基準	検討事項
Intellectual merit (知的価値)	<ul style="list-style-type: none"> ・ 提案案件の知的価値は何か。 ・ 分野内や他分野における知識・理解の進展に対し、どの程度の重要性を持つか。 ・ 提案者（個人または団体）は、プロジェクト遂行において適切な資質があるか。（該当する場合、評価者は先行研究の質についてコメント。） ・ 提案案件は、創造的、独創的、または変化をもたらしうる考えをどの程度示唆し、探求しているか。 ・ 提案案件は、構想・構成が優れているか。 ・ リソースへのアクセスは十分あるか。
Broader impacts (波及効果)	<ul style="list-style-type: none"> ・ 提案案件の波及効果には何があるか。 ・ 教育・訓練・学習の促進ならびに、発見・理解の進展がどの程度もたらされるか。 ・ 現時点では参画が不十分なグループ（例：性別、民族、障害、地域等に偏りがある）の参加を拡大できるか。 ・ 施設、設備、ネットワーク、パートナーシップ等、研究・教育の基盤をどの程度強化できるか。 ・ 成果が広範に発信され、科学技術に関する理解が強化されるか。 ・ 提案案件が社会にもたらしうる利点には何があるか。

(出典) NSF Grant Proposal Guide Chapter III <http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_3.jsp> [last accessed: 2012/2/3]

なお、「broader impacts」については、判断が困難なことも想定されるとして、事例を示した文書⁽⁸⁾が発行されている。この概念は、当初科学者コミュニティにはなかなか受け入れられず、提案内容・選考での議論に反映することが困難だったとされている⁽⁹⁾。NSFがこれら2つの基準について記載していないと申請を返却するとしたのは2002年になってから⁽¹⁰⁾であり、返却される申請は、2003年には276件あったが、減少傾向にあり、2008年には124件である。

NSFの活動を評価する外部の専門家で構成された委員会である外部者委員会（Committee of Visitors: COV）が2004年から2007年にかけて行ったNSFに対する評価では、評価プロセスに関する問題点が指摘されている⁽¹¹⁾。特に「broader impacts」という基準については、定義が曖昧、評価における詳細不足、基準の重要性に関する認識不足、といった問題が指摘されている。

NSFの方針を決定する国立科学理事会（National Science Board: NSB）は2010年2月にメリット・レビューに関するタスクフォースを設置し、2011年1月にメリット・レビューに関する意見を招請し、6月には案を公表してさらに意見募集を行った。新たな基準案では、broader impactsについて、どのような国家目標に貢献し、それが適切な体制、資源、アプローチで行われる合理的な計画となっているかを問うようになっている⁽¹²⁾。2011年12月には“National

(8) “Merit Review Broader Impacts Criterion: Representative Activities”

<<http://www.nsf.gov/pubs/gpg/broaderimpacts.pdf>> [last accessed: 2011/12/09]

(9) 独立行政法人科学技術振興機構『海外の競争的研究資金配分機関における研修、および競争的研究資金制度調査』2006, p.60.

(10) “Science for the masses” <<http://www.nature.com/news/2010/100526/pdf/465416a.pdf>> [last accessed: 2011/12/09]

(11) “Report to the National Science Board on the National Science Foundation’s Merit Review Process, Fiscal Year 2008” p.31. <<http://www.nsf.gov/nsb/publications/landing/nsb0943.jsp>> [last accessed: 2011/12/09]

(12) National Science Board – Publications <http://www.nsf.gov/nsb/publications/2011/06_mrtf.jsp> [last accessed:

Science Foundation's MERIT REVIEW CRITERIA, REVIEW AND REVISIONS”として調査結果、検討結果がまとめられた⁽¹³⁾。この報告書では、broader impactsについてはintellectual meritに比較して研究者コミュニティやNSFスタッフに理解されていないという問題点はあるものの、これら2つの基準は維持するものとし、原則（principles）を追加するとともに、基準同士の関係を整理することを提言している。broader impactsについては、研究の多様性を考慮して、特定の活動を推奨することはしないようにしている。

(2) NIH (National Institutes of Health)

NIHは、自身が医学研究機関であると同時に大学等の研究機関に資金配分を行っているファンディング・エージェンシーである。その競争的グラントの中核となるものがR01と呼ばれる研究プロジェクトグラントプログラムであり、研究者のボトムアップによる申請に基づき競争的に配分されるグラントで支援機関は3～5年間である。

NIHにおいては他にも様々なプログラムが実施されているが、R01は従来の枠を外れた革新的な発想に対する支援を行うためには必ずしも適当な制度ではないという、ピア・レビューにおける保守的な側面に対する指摘を受け、2004年にNIH所長の主導によりNIH内の研究所横断型プログラムであるパイオニア・アワード（NIH Director's Pioneer Award Program (NDPA)）が創設された。NDPAはNIHロードマップの一環として設置された、いわゆるハイリスク研究支援プログラムで、少数の極めて高い創造性を持つ研究者個人を支援することを目的としている。

NDPAの評価手順・方法は、R01をはじめとする他のNIHのプログラムと大きな違いがある。まず、申請書類のうち、研究計画に関する部分は5ページまでのエッセイと1ページの申請者の顕著な研究業績により構成されており、研究計画だけでも12ページの研究戦略（research strategy）を求めるR01より少なく⁽¹⁴⁾、研究内容の記載について高い具体性は求められていない。

また、R01の書面と会合によるピア・レビューとは異なり、外部の研究者によるシンプルなオンラインレビューと面接の2段階で行われる。ピア・レビューの第1段階では外部の研究者によるオンラインレビューが行われ、特に高い評価を得た申請のみに対し、第2段階として卓越した業績を持つ研究者のパネルによる面接審査が行われる。最終的な決定はNIH所長の諮問委員会の議を経て、NIH所長が行う。

ピア・レビューにおける評価は、他のプログラムと同様の基準と同時に、申請において提示された科学的課題やチャレンジ、研究者（研究者の革新性・創造性など）、NDPAによる支援の趣旨との適合性（リスクが高く見返りが大きい計画か）といった観点において行われる⁽¹⁵⁾。

2011/12/09]

(13) "National Science Foundation's MERIT REVIEW CRITERIA, REVIEW AND REVISIONS" <<http://www.nsf.gov/nsb/publications/2011/meritreviewcriteria.pdf>> [last accessed: 2012/2/7]

(14) 研究戦略（research strategy）は2009年まで25ページであった。

(15) NIH Directors Pioneer Award Program (DP1) Section V. Application Review Information, Review and Selection Process <<http://grants.nih.gov/grants/guide/rfa-files/RFA-RM-09-010.html#SectionV>>, [last accessed: 2012/2/3]

3 イギリス

大学に対しては機関単位の一括交付金に加えて、分野ごとのファンディング機関である研究会議 (Research Councils: RCs) が競争的に配分しており、この点では日本と類似のしくみとなっている。

(1) BBSRC (Biotechnology and Biological Sciences Research Council) のsLoLas

RCsの1つである英国バイオテクノロジー・生物科学研究会議 (BBSRC) でも、通常のファンディングとは別にStrategic longer and larger grants (sLoLas) と呼ばれる挑戦的な課題を対象とした制度が設けられている⁽¹⁶⁾。

この制度は2006年にLoLasとして挑戦的で、学際的な研究提案を支援するために設けられたものであり、2009年にはよりBBSRCとそのステークホルダーの戦略的な重要領域に重点が置けるように修整が行われている。

具体的には、sLoLasの提案は以下の2つを満たす必要があり、大型で長期の研究を支援するように設計されている。

- 200万ポンド超の予算であること
- 5年までの期間であること

審査基準としては以下が特徴的である。

- 科学的に卓越していること
- 1つ以上のBBSRCの戦略目標に合致していること
- インパクトが示せること
- 国際的に先導的な研究チームで実施されること

III まとめ

研究開発はそれぞれ多様な目的や期待される成果をもっており、その性格に応じて評価方法や観点も自ら異なってくる。研究開発評価は一面的・画一的に実施されると、それに対して過剰な適応が行われるために、かえって研究開発の目的の達成が阻害される場合がある。例えば、短期的な成果を過度に求めるために、人材育成など中長期的に研究基盤を維持するための活動が損なわれることや、失敗のリスクが少ない小粒なテーマに偏重したりすることが挙げられる。

我が国では政策評価法に基づく「国の研究開発評価に関する大綱的指針」で評価の観点として必要性、効率性、有効性を挙げている。各府省による研究開発評価もこれを基本としているが、研究開発の性格や目的による違いが見られる。たとえば、学術研究を支援する制度である科学研究費補助金の審査では、研究課題の学術的重要性、独創性、波及効果や研究遂行能力を問うているが、研究開発課題の政策課題との合致や研究開発成果の確実性は直接問うていな

(16) “Strategic longer and larger grants (sLoLas)” <<http://www.bbsrc.ac.uk/funding/grants/lola/lola-index.aspx>> [last accessed: 2011/12/09]

い。一方、科学技術振興機構の戦略的創造研究推進事業では、文部科学省が示した戦略目標の達成への貢献が評価基準に挙げられている。

米国では、研究費の配分は競争的な方法が主であるため、その中に多様な評価の視点を設ける試みがなされている。医学を除くあらゆる基礎科学・工学を支援しているNSFの審査方法はメリット・レビューとして知られており、Intellectual meritとBroader impactsの2つを評価している。このうちBroader impactsについては、当初からなかなか科学コミュニティに受け入れられず、現在でも定義が曖昧、評価における詳細不足、基準の重要性に関する認識不足、といった問題が指摘されて、見直しが続けられている。

ハイリスクの研究については、我が国ではさきがけの「大挑戦型」として実現可能性が得難いものを評価する視点を持つ制度が追加的に設けられている。米国のNIHでも、中核であるR01とは別に、パイオニア・アワード（NDPA）と呼ばれるハイリスク研究支援プログラムがある。イギリスのファンディング機関の1つであるBBRSCでも、通常ファンディングとは別にStrategic longer and larger grants（sLoLas）と呼ばれる挑戦的な課題を対象とした制度を準備している。

このように、各国の事例においても、研究開発の多様性や不確実性に配慮した特徴ある評価基準が設けられているが、望ましい手法については試行錯誤が行われている。