

教育行政領域における実証的研究の権威性の正当性に関する研究

—No Child Left Behind Act of 2001 を事例とした第一次的検討—

桐村 豪文

はじめに

近年、実証的研究という言葉がもてはやされている感がある。その言葉は多面でもてはやされ、用いられたため、その言葉について共有しうる基底の意味が何か、不明確になっている感も否めない。事実、ウィーン大学に発祥の地をもつ「論理実証主義」、あるいはその強い影響を受けて1950年代から1960年代初頭にかけてアメリカにて理論運動として「展開され、教育行政事象を自然科学的な方法で捉えようとする実証主義パラダイム」¹で理想とされた理論は、仮説—演繹理論であり、「理論を純粋に論理—数学的手続きでもって、より多くの一組の経験的法則を導き出すことのできる一組の過程として定義」²されるものであった。今日実証的研究と自称する研究は、例えば歴史研究など、原初の「実証的研究」がもつ意味の埒外で用いられる場合があり、したがってその言葉について共有される意味がやや多面性を持ってきたと言える。しかし、高木太郎（1967）が「教育行政の実証的研究」を「社会学的研究と統計的研究の二つ」³に区分していることや、中澤渉（2003）⁴によるこれまでの実証的研究に対する自省的考察などを踏まえると、全体的には従前の論理—数学的な形態、方法論を維持していると言えよう。

しかし本研究では、実証的研究という言葉について、共有すべき基底の意味を、一連の自称実証的研究の事例の中から剔出することを目的としているのではない。この言葉が持つ問題性は曖昧性にあるのではない。その曖昧性を生成するに至らしめた権威性にあるのである。権威性を可能とする言葉であるがゆえに、多面でもてはやされ用いられたのである。これは、かつてブルデューが「ロジ—効果」と呼称した、「科学と哲学の間の境界を混乱させようとする哲学者たちの努力の明らかな表現」⁵に類似するものである。実証的研究という言葉も同様、研究に対する権威付けを曖昧にも保証する力をもつ言葉となっているのである。

ではなぜその権威性が問題となるのか。より厳密に言うならば、権威性の在り方に問題を見る。語りが権威性を持つということ全てが問題となるのではない。例えばトートロジーとして語りうることは全て権威性をもつ。なぜならば、語りが権威性をもつのは、真理性をもつからである。つまり、トートロジーは常に真であるから、何人もこれを偽とする自由は得ておらず、その意味においてトートロジーは常に権威性を持つのである。

よって、ある語りが権威性を持つ根拠は、その真理性に拠るのである。この論理的順序は定言的事実である。ただしこの場合、語りが真理であることは、語りが権威性を持つための十分条件としてあるが、必要条件ではない。たとえ語りが真理でなくとも、例えば、語りの主体が権威者（政府など）であることや、あるいは方法論が語りに対して隠蔽的に真理性を帯びさせることを

可能とさせる、といったことから、語りに権威性が与えられる事態は発生しうるのである。この事態はつまり、語りに対して、恣意性がその内に隠匿された権威性が、あるいは語りうる真理の土台を等閑に付し、自らを真理性の限りに制約しないことを、真理性とは別の力によって隠蔽することによって得られる権威性が与えられるということを意味するのである。ただし本研究では後者の問題（方法論）に着眼したい。前者の問題（語りの主体）は、権威性の度合いに対して直接作用する要素であるが、真理性を隠蔽する論理構造をもったものではない。対して、方法論は意図的如何に関わらず、結果としてそうした機能を果たしうる力をもっているのである。

ある語りを成立させるために依拠する方法論は、本来それ自体真理に基づかなければならないが、しかし真理性を再検討されることなく、真理性ではなく、一定の信頼性の上に活用され続けることがある。本研究で取り上げる実証的研究の場合もそうであり、一定の方法論を遵守することで、その方法論が真理性という面ではいかなる位置づけにあるかを再検討することなく、信頼性の下、自動的に権威性が保証されるのである。よって権威性の在り方について問題であるのは、真理性の検討を待たずして、権威性の与えられた方法論に依拠することで、あたかもその語りが真であるかのように振る舞い、結果として権威性を獲得し行使する、ということである。

またこの権威性は、研究の領域に留まるものではない。例えば、実証的研究に基づくことが前提とされるような教育行政や、あるいはそうした行政作用の下に置かれる教育実践も、その影響の及ぶうる範囲の内にあるのである。すなわち、教育行政活動、教育実践活動が遂行される上で実証的研究に基づくことが求められるということは、それら諸活動が、実証的研究において真理とされる限りに制約されるということの意味するのである。ここで問題となりうるのは、諸活動が制約されるという場面における、真理性としての基準（方法論）の定められ方においてである。つまり、制約されるという事態それ自体が、実践者に対する桎梏として問題視されるのではなく、実証的研究における真理性の基準の定められ方に、先述の問題を導出する契機となる可能性が潜んでいるのである。もしこの真理性の基準の定められ方に恣意性が見出されるならば、正当性を有さない権威性の現出と見うるのである。よって、実証的研究のもつ権威性の正当性を検討する上では、真理性の視点から、その隠匿された恣意性を基準の内から剔出しなければならない。

以上に示すとおり、本研究の究極的な目的は、教育行政領域における実証的研究について、その「実証的である」とする基準（方法論）の正当性に対して、真理性の視点から改めて検討を加え、結果として、実証的研究が持ちうる権威性の正当性を検討することにある。しかしこの問題は非常に困難なものである。基準（方法論）の正当性を検討するためには、前提として、真理性とは何か、このアポリアに一定の決着をつけないければならないのである。真理として語りうるための条件は何か。さまざまに規定され、制約される現存在として、真理として語りうるものは何か。この問題は、哲学的考究を通してしか決着を見ない。しかし本研究の究極的目的を達成するための最も重要な必要条件がここにあることは確かである。

真理性に関する哲学的考究という任務が、本研究の究極的目的を達成する上での必要条件として示されたわけだが、もう1点、まず果たすべき、目的達成上必要な任務がある。それは、検討されるべき対象事例から、規定される「実証性」の基準を示すことである。基準の正当性あるいは恣意性を、規定される基準の内から剔出するためには、まずそれら基準を、自ら顕すとおりに記述する必要がある。したがって、もし示される基準が複数存在するならば、それら諸基準の相

対的位置付けも併せ考察することが、諸基準の真理的構造の顕現化にとって有益であると考えられる。

本稿では、本研究の第一次的段階として、この後者の任務を果たすことを目的とし、米国の No Child Left Behind Act of 2001 (以下、NCLB 法) を検討対象事例として取り上げる。同法は、教育行政活動、教育実践活動の裏づけとして法的に実証的研究の組み込みを求めており、したがって本研究の目的に照らしたとき、事例として取り上げるに相応しいと考える。なお、以下実証的研究と表記するものは全て、同法に示される “scientifically based research” を意味する。

1. NCLB 法における実証的研究の位置づけ

NCLB 法は、選挙中から教育を最優先事項として掲げていた G. W. ブッシュ (George W. Bush) 大統領が、就任して 2 日目に提出した初の法案であり、その後 2002 年 1 月 8 日にブッシュ大統領が署名し、成立した。NCLB 法は、初等中等教育法 (Elementary and Secondary Education Act) の修正法である。NCLB 法には 4 つの柱 (pillars) が立てられている。それは、①結果に対するより強固なアカウントビリティ (Stronger Accountability for Results)、②州や地域に対する裁量権の拡大 (More Freedom for States and Communities)、③実証された方法 (Proven Methods)、④保護者に対する選択権の拡大 (More Choices for Parents) である。本稿で取り上げる実証的研究は、その 4 つの柱の内の「③実証された方法」に関わるものである。

NCLB 法の中で実証的研究という言葉は、大きいプログラムから小さいプログラムに至るまで、計 111 回出現する。よって NCLB 法では、政策的にその重要性が提唱されているのみならず、法的に教育行政活動及び教育実践活動の中に実証的研究が組み込まれているのである。したがって例えば、教育実践者は、直観ではなくエビデンスに基づくことが求められ、また連邦補助金の受給の可否は、プログラムが実証的研究に基づいているか否かによって決定されるのである。

本稿では以下、初等中等教育法に準拠する連邦補助金プログラムの内、実証的研究に関連する事例として、「読解力向上プログラム (Reading First)」、「包括的学校改善プログラム (Comprehensive School Reform)」の 2 つの補助金プログラムを取り上げる⁶。取り上げる理由は、両者が共に実証的研究の定義について、一般的定義とは別に独自に定義を設けているためである⁷。以下、この両プログラムに関して、まず 1.1 と 1.2 において、そのプログラムの概略と実証的研究との接点を素描する。そして 2. で、一般的定義と各プログラムの定める実証的研究の定義を示し、各定義の特徴を、3 者間の比較考察を通じて、顕現化する。

1.1. Reading First (Title I, Part B, Subpart 1)

Reading First について概容を示すと、以下のとおりである。このプログラムは、Reading Excellence Act に代わり、NCLB 法 Title I, Part B, Subpart 1 によって創設されたものである。連邦教育省初等中等教育局の学力・教員改善課 (Academic Improvement and Teacher Quality Programs) の中の、Reading First の部門が管理している。2008 年度の補助金額は 3 億 9301 万 2 千ドルであり⁸、その補助方式は一般補助方式である。ただ、州に対する配分方式は一般補助方式だが、州から地方教育当局に対して再配分される場合は競争的資金として配分される。州と学区を受給対象者とし、実証的研究に基づき、就学前から第 3 学年までを対象に読解 (reading) を教授するプログラムに対して支給される。Reading First の目的は、読解に関する実証的研究

に基づいた教科指導プログラムや教科指導上の資源、成績評価、職能開発を実行することを通じて、第3学年の終わりまでに、全ての児童が学年に応じた、またはそれ以上の読解能力を確実に修得できるようにすることにある。

このプログラムの場合、その目的の記述の内からすでに実証的研究への言及がなされているので、とりわけ実証的研究との密接な関係が前提としてあることがわかる。このプログラムにおける実証的研究の位置づけに関してさらに言うと、以下のように規定されている。

- ① 州教育当局または学区は、実証的研究が読解の教育指導上必須の要素としてみなす5つのキーエリア（音素の認識、音響学、語彙、流暢さ、読解の理解）に焦点を当て、職能開発や教科指導上のプログラム、資源を選択しなければならない
- ② Reading First は、教員が実証的研究に基づく教育指導プログラムを教授することが可能となるよう、研修を増やし、提供する

1.2. Comprehensive School Reform (Title I, Part F)

次に Comprehensive School Reform (CSR) だが、このプログラムは、1998年に創設された Comprehensive School Reform Demonstration Program (CSR/D) に代わり、NCLB 法 Title I, Part F に準拠し、Comprehensive School Reform に名称変更されたプログラムである⁹。Reading First と同じく、このプログラムを管理する部門は、連邦教育省初等中等教育局学力・教員改善課の中にあり、CSR のみを担当する部門がこれを管理している。2008年度の補助金額は160万5454ドルで¹⁰、Reading First 同様、補助方式は州に対しては一般補助だが、地方学区に対する再配分においては競争的配分方法がとられている。このプログラムの目的は、その有効性が実証された手法、方策を用いることによって、包括的な学校改善を施し、児童生徒の学力を向上させることを目的としている。CSR とその前身の CSR/D とを比較するとき、プログラムが満たすべき構成要素において違いが見られる。CSR/D では、補助金を獲得するためにプログラムが満たさなければならない構成要素が9つであったのが、CSR ではそれに、下記の①と⑥の2つの要素が加えられ、以下の11の構成要素から成っている。

- ① 教育、学習、学校運営において、実証的研究や有効な実践事例に基づいた方策、方法の採用
- ② 一連の構成要素と包括的設計図とを統合させること
- ③ 教職員の、継続した質の高い職能開発の提供
- ④ 児童生徒の学力を対象とした評価可能な目標とベンチマークを含むこと
- ⑤ 学校内において、教員、行政官、職員によって支援が行われること
- ⑥ 教員、行政官、職員に対して支援を提供すること
- ⑦ 保護者、地域が学校の改善活動の計画、実施、評価へ有意義に参加することに備えること
- ⑧ 学校全体の改善・改革に長けた外部協力者から、質の高い技術支援、補助を取り入れること
- ⑨ 学校改善の実施計画を策定すること、及び毎年達成されるべき児童生徒の学力成果に基づく方策を策定すること、さらにその方策に対する評価についての計画を策定すること
- ⑩ 学校の包括的改善努力を支援、支持するための資源を確保すること
- ⑪ 実証的研究を通じて、児童生徒の学力向上をもたらすことが証明されていること、もしくは、学力向上をもたらすことを示す強固なエビデンス (strong evidence) をもつこと

構成要素⑩に示されるように、CSR は、前身の CSR D の時期から、より科学的手法に基づくことが前提とされたプログラムであった。①と⑩のダブルスタンダードの問題については、2.2. で後述するが、NCLB 法は①を追加したことで、この科学性の条件をより強固なものとしたのである。つまり、補助金を受給するためには、プログラムが、その有効性が実証的に示された改善モデルを選択しなければならないということを義務付けたのである。

また CSR には、CSR プログラムの質を高める活動を支援するため、CSR イニシアティブ (Comprehensive School Reform Quality Initiatives) というプログラムが並立され、非営利組織などに対して補助金を提供してきた¹¹。連邦教育省初等中等教育局は、ジョンズ・ホプキンス大学 (Johns Hopkins University)、アメリカ調査研究所 (American Institutes for Research) などと提携を結び、技術支援、あるいは学校改善モデルの構築を求めた。

2. NCLB 法における実証的研究の定義

1. で素描した、実証的研究との接点を持つ 2 つの連邦補助金プログラムについて、その定める実証的研究に関する定義、実証性の基準 (方法論) を、その隠匿する恣意性まで含め、顕現化させることを、本節では目的とする。まず、2.1. で NCLB 法第 9101 条において定められる一般的定義を、ヴァレリ・レイナの解説を踏まえ、示す。続いて 2.2. と 2.3. にて、CSR と Reading First における、一般的定義とは若干立場を異にする定義を、各プログラムが頒布するガイダンスから引用し、示す¹²。図表 2、図表 3 は、ガイダンスにおける実証的研究の規定に関する内容の一部である。そこでは具体的問いの形を用いて、実証的研究の基準がより詳細に示されている。そして 2.4. で、それら 3 者間の比較考察から、隠匿されたる恣意性の顕現化を試みる。

2.1. 一般的定義

NCLB 法第 9101 条の下、全てのプログラムを対象とする実証的研究の定義が与えられている (図表 1)。この定義 (A) によると、「厳格で体系的で客観的な手順」が、実証性の根底的基準として設定されている。では「厳格で体系的で客観的な手順」とは具体的に何を指しているのか。

この定義 (B)(iv)によれば、研究の方法論として実験的手法を用いることが、最も強固な厳格性、体系性、客観性をもった分析手法であると言う。ここで言う実験的手法は、無作為統制 (randomized controlled) 実験とも呼ばれ、最も優れた方法論として「golden standard」とまで称される¹³。この方法論は、医療や福祉、雇用政策、心理学など他の分野で、あるモデルプランの有効性を評価するときに、最も優れた手法として活用されてきたのだと言う。

ではさらにここで言う無作為統制実験とは何か。これについては、当時教育調査改善局 (Office of Educational Research and Improvement) ¹⁴次長 (Assistant Secretary) の上級アドバイザー (senior research advisor) であったヴァレリ・レイナ (Valerie Reyna) ¹⁵が、連邦教育省が 2002 年 2 月 6 日に開催したワーキンググループ会議で行った説明を援用して解釈してみよう¹⁶。それによると、無作為統制実験とは、「例えば、あなたが、無作為統制実験で、新しい第 3 学年の数学のカリキュラムが現行のカリキュラムよりもより効果的か否かを検証したいと望んでいるとする。あなたは、多くの第 3 学年の児童を対象に、新しいカリキュラムを活用する介入グループ (intervention group) か、現行のカリキュラムを活用する統制グループ (control group) の

図表 1 NCLB 法における、実証的研究についての定義

- 「実証的研究」の用語は、
- (A) 教育活動やプログラムに関する信頼ある、確かな知識を獲得するため、厳格で体系的で客観的な手順を持った真理性を含む研究を意味する。
 - (B) そして以下のような研究を含む
 - (i) 観察や実験に基づいて設計する体系的で実証的な手法を用いること
 - (ii) 仮説検証から導出された一般的結論が正しいことを証明するために十分かつ厳密なデータ分析を含むこと
 - (iii) 全ての評価者や観察者において、全ての多様な評価・観察において、全ての様々な調査者による研究において、信頼ある、確かなデータを提供する観察手法、評価手法に依拠すること
 - (iv) 個人、団体、プログラム、実践に異なる条件が与えられ、重要な条件の影響を評価するために適切な統制を用いて、無作為に介入する実験を優先させる、実験的または準実験的計画を用いて、あるいは、もし実験的・準実験的計画が条件の内外に統制を含むならば、他の計画を用いて評価されること
 - (v) 確実に、実験的研究が、再現を考慮するに十分詳細で明快に提供されること、あるいは少なくとも、得られた知見の下で体系的に研究を行う機会を提供するために十分詳細で明快に提供されること
 - (vi) 相対的に厳格で客観的で科学的な検討を通し、他者評価（ピアレビュー）を受けた雑誌で受け入れられること、または独立した専門家委員会によって認められること

【出典】The No Child Left Behind Act, Section 9101.

いずれかを無作為に割り当てる。あなたはそうして時間が経過した後、両方のグループの数学の成績を評価する。その2つのグループ間の数学の成績の違いは、現行のカリキュラムに対する新しいカリキュラムがもつ効果を表している」¹⁷と云う。つまり、ここで言う無作為統制実験とは、すでにあるサンプルを介入グループ、統制グループに無作為に割り当て、その両者間で現れる効果の違いを比較観察するという対照実験を言うのである。

そしてさらに彼女いわく、この無作為統制実験は、他の方法論にはない、独特の利点があるのだと云う。つまりそれは、「多くの個人を介入グループまたは統制グループに無作為に割り当てるプロセスによって、どんな特性（観察されるものも観察されないものも）においても、グループ間には体系的な違いは存在しないということが保証され」、結果として、「介入グループと統制グループとの間で生じた結果の違いは、信頼性を以って介入に起因しており、他の要因には起因していないといいうる」¹⁸という特長である。つまり正真正銘、対照実験として、現れる効果の違いを、操作を加えた特定の変数のみに起因するものと断定できると云うのである。

またこの定義に示されている「厳格で体系的で客観的な手順」は無作為統制実験のみではない。定義(B)(iv)に示されているように、実証的研究に認められるものとして、準実験的(quasi-experimental)研究も含まれている。準実験的研究とは、比較グループ(comparison group)研究とも呼ばれ、無作為とは別の手法を通じて選ばれた比較グループと介入参加者とを対比させ、その両者間に見られる結果を比較するものである。

連邦教育省が発行した実証的研究に関するガイドによれば、準実験的研究は、「よく調整された研究は、医学やその他分野と同様、教育の分野でも、介入の有効性について『真でありうる(possible)』エビデンスを打ち立て、無作為統制実験によって確証を得るに値する仮説を構築するという点で、価値ある役割を果たしうると考える。しかしそのエビデンスは、十分に調整された比較グループ研究でさえ、何が効果的か否かの最後の裁定者として、あるいは効果の強さに対する信頼あるガイドとして活用されることに対しては、注意する必要がある」¹⁹と。このように準実験的研究について連邦教育省は、その有効性については限定的に捉えているが、しかし実験

的研究に比類するものとして価値づけている。

2.2. Comprehensive School Reform における実証的研究の基準

CSR では、1.2. で示したように、実証性という点に関して、2 つの基準が設定されている。すなわちそれは、(1) 実証的研究に基づくということと、(2) 強固なエビデンスをもつということである。だが、両者は相矛盾するものではない。この「強固なエビデンス」は、実証的研究に準ずる概念なのである。では、ここで言う「強固なエビデンス」とは何か。

上述の CSR のガイダンス²⁰は、「強固なエビデンス」という基準について、以下のように説明している。『「強固なエビデンス」とは、実証的研究に基づいたエビデンスよりも、より厳密さを欠く基準として定義される。『強固なエビデンス』として積極的効果を示す実践や方策、プログラムは、実証的研究の定義で設定された基準を満たす、広範な研究基盤を欠いている』²¹と。この説明から明らかな点は、「強固なエビデンス」は「広範な研究基盤」を欠き、実証的研究に基づいたエビデンスよりも厳密さを欠いたものであるということである。しかしいまだ不明な点は「広範な研究基盤」という言葉である。それは以下のような構造を意味している。

まず CSR においては、図表 2 に示す 6 つの基準全てを満たす事例研究を「質の高い事例研究 (High-quality study)」、6 つのうち 5 つを満たすものを「穏当な事例研究 (Reasonable quality study)」と呼んでいる。実証的研究は、少なくともこの「質の高い事例研究」5 つから構成されていることが必要条件として課されている。他方その条件を満たさず、「質の高い事例研究」と「穏当な事例研究」で複合的に構成されている研究を「強固なエビデンス」と称するのである。

1.3. で示した構成要素⑩からわかるとおり、CSR では、実証的研究という基準のみならず、以上に示した、それよりもやや厳密さを欠く「強固なエビデンス」という基準も、同時に認められている。しかし、厳密さに欠けるけれども、ガイダンスは以下のようにも述べる。つまり、強固なエビデンスは、これらの基準の全てを満たすものではないけれども、なお依然として科学性を基盤とするものではある、と。これは、図表 4 で示されているように、そもそも CSR が他の 2 つの定義よりも実証的研究に対してより厳密な科学性を要請しているから、ということもあろう。

2.3. Reading First における実証的研究の基準²²

連邦教育省が発行してきた、実証的研究に関するガイダンスは、そのほとんどが Reading First を対象に作られたものであった。それにはいくつかの理由があるが、主として、その立ち上げの経緯に起因する。Reading First はそもそも、新しいプログラムとして、実証的研究に基づいた読解アプローチを重要視するという考えから始まった。またこのプログラムの創設者スーザン・ノイマン (Susan Neuman) は初等中等教育局次長 (Assistant Secretary for Elementary and Secondary Education) を務め、また彼女自身読解に関する研究者でもあった。こうした経緯から、NCLB 法の他のプログラムとは異なり、Reading First はとりわけ実証的研究と密接不可分の関係にあるのである。

図表 3 に示されるように、Reading First に適用される実証的研究の定義は、NCLB 法に適用される一般的定義と大きく異なる。例えば、Reading First の定義では、研究の方法論が実験または準実験であることを求めない。基準を緩和することで、より多くのプログラムとアプローチ

<p>図表 2 実証的研究か否か評価するときに用いられる問い CSR ガイダンス (B-3)</p> <ol style="list-style-type: none"> 体系的で帰納的な方法 研究は確固たる理論的土台を持っているか。 方法論、被験者、研究者は明確に示されているか。 研究は、矛盾のない、統制のとれた、方法論的手法をもって実行されたか。 データは、観察や実験を用いて得られたか。 研究は、主観的考えよりも事実即ちデータに基づいたものか。 研究で得られた知見は、具体的で、知覚可能なエビデンスによって支えられたものか。 厳密なデータによる分析 研究は、提示された仮説を検証し、一般的結論を導き出したか。 研究は、サンプルのサイズや用いた統計手順を報告しているか。 データは、任務に適した手法を用いて分析されたか。 手法は、データの特徴や構造に合致しているか。 研究は、観察された結果に対して他の説明が生じないように設計されているか。 研究で得られた知見は、観察された結果が、特定の介入が原因であると、納得させる証拠文書を提示したか。 信頼できる、妥当なデータの収集 データは、多くの場で多元的な調査者が関与した研究から導き出されたものか。 研究のバイアスは最小化されたか。 データは矛盾なく測定されたか。同じ条件の下で再度行われた測定では、同様の結果を導き出したか。 強固な研究計画 研究設計は、適切な統制の下、被験者を異なる条件に無作為に割り当てる、無作為統制実験を表しているか。 統制は、関心のある条件の評価を可能としているか。 研究は、調査者が自身のリサーチクエッションに解答できるよう、最も効率よく設計されたか。 再現を可能とする詳細な結果 得られた知見は、明確に記述され報告されているか。 研究の結果は、設計の再現が可能となるよう、十分詳細に示されているか。 得られた知見は、今後の研究と共に、その可能性を高めうるか。 調査を受けた結果 研究は、ピアレビューが行われる雑誌あるいは独立した専門家によって認められたものか。 研究は、研究活動の一員ではない、バイアスのかかっていない専門家によって検討されたものか。 レビューを行う者は、厳格な学問的水準を適応し、質の高い統制を提供したか。 研究は、外部の検証を受けたか。 <p>(出典) U.S. Department of Education, Comprehensive School Reform (CSR) Program Guidance, 2002</p>

<p>図表 3 実証的研究か否か評価するときに用いられる問い Reading Firsts ガイダンス (B-3)</p> <ol style="list-style-type: none"> 厳密で体系的で経験的な方法の使用 研究は、確固たる理論的基盤、あるいは研究的基盤を持っているか。 研究は、バイアスのかかった発見や不当な有効性の主張を避けるため、注意深く設計されたものか。 研究は、どのように行われ、誰によって行われ、何について行われたか、明白に記述しているか。 提示された仮説を検証し、一般的結論を正当化するためのデータ分析の適切性 研究は、観察された結果に対して他の説明が生じないように設計されているか。 観察された結果は、有効性についての全体的結論や主張と矛盾していないか。 研究は、観察された結果が特定の介入の結果であると、納得させる証拠文書を提示しているか。 研究は、どういった人が研究されたか(被験者の年齢、地理的、認知的、学力的、行動的特徴)について明確に示しているか、そして得られた知見はどういった人に対して一般化しうるのか示されているか。 研究は、成果に対する評価について十分な記述を提示しているか。 評価者や観察者、多元的評価や観察に関わる妥当なデータを提供する評価の手法、観察の手法への依拠 データは、一人の調査者、一つの授業に基づいたものか、もしくは多数の場所における多数の調査者によって集められたデータか。 研究者のバイアスを最小化するためにどういった手順が実行されたか。 観察された結果は、時間が経過しても有効であるか。 研究対象の介入は、再現可能なものとして、詳細なデータをもって示されているか。 研究は、教育指導上の迫真性がどのように確保され、評価されているか説明しているか。 査読雑誌による受容、あるいは比較的厳密で客観的で科学的な再検討を通じた、独立した専門家による認可 研究は、研究活動の一員ではない、バイアスのかかっていない個人によって慎重に検討されたか。 得られた知見は、外部の綿密な調査と検証を受けたか。 <p>(出典) U.S. Department of Education, Guidance for the Reading First Program, 2002</p>
--

図表 4 3つの定義の相互比較	一般的 定義	CSR	Reading First
■ 理論性、研究設計			
体系的であること	○	○	○
多面的評価や観察を用いること	○	○	○
観察や実験を用いること	○	○	
確固たる理論的土台を持っていること		○	○
方法論、被験者、研究者を明確に示すこと		○	○
観察された結果に対して他の説明が生じないように設計すること		○	○
バイアスを最小化すること		○	○
矛盾のない、統制のとれた、方法論的手法を用いること		○	
自身のリサーチクエッションに解答できるよう、最も効率よく設計すること		○	
関心のある条件の評価を可能とするよう統制を行うこと		○	
■ 分析手法 データ			
仮説検証型	○	○	○
厳密なデータによる分析を実行すること	○	○	○
一般的結論を得ることを目的とする	○	○	○
適切な統制を用いて、無作為に介入する実験を用いること	○	○	
主観的考えよりも事実即したデータを用いること		○	
知覚可能なエビデンスを用いること		○	
任務に適した手法を用いること		○	
データの特徴や構造に合致した手法を用いること		○	
データは矛盾なく測定されていること		○	
観察された結果と全体的結論との間に矛盾がないこと			○
■ 成果報告 知見の応用			
サンプルのサイズや用いた統計的手順を報告すること		○	
観察された結果が、特定の介入が原因であると、納得させる証拠を提示すること		○	
観察された結果が特定の介入の結果であると、納得させる証拠を提示すること			○
被験者の年齢、地理的、認知的、学力的、行動的特徴を示すこと			○
得られた知見はどういった人に対して一般化しうるのか示すこと			○
成果に対する評価を行うこと			○
妥当なデータを提供する評価的手法、観察的手法へ依拠すること			○
観察された結果が、時間が経過しても有効であること			○
教育指導上の迫真性が確保されていること			○
同条件の下で同様の結果が導き出され、再現可能であること	○	○	○
再現を可能とする詳細な結果を提示すること	○	○	○
■ レビュー			
他者評価（ピアレビュー）を受けた雑誌によって認められること	○	○	○
研究員ではない、バイアスのかかっていない専門家による検討を受けること	○	○	○
外部の検証を受けること		○	○
レビューを行う者は、厳格な学問的水準を適応し、質の高い統制を提供すること		○	

が資格あるものとして認められるのである。

このように、Reading First は一般的定義と大きく内容を異にしている。それは上述のとおり、当初から実証的研究の活用に焦点が当てられていたため、一般的定義によって焦点の変化を求めする必要はなかったということが背後にあるからである。

2.4. 小括

本項では、以上3つの定義内容、かつ図表4に示した三者比較を踏まえ、実証性の基準から恣意性の顕現化を行いたい。

まず3つの定義に共通するものとして、「体系的であること」「多元的評価や観察を用いること」「仮説検証型」「厳密なデータによる分析を実行すること」「一般的結論を得ることを目的とする」「同条件の下で同様の結果が導き出され、再現可能であること」「再現を可能とする詳細な結果を提示すること」「他者評価（ピアレビュー）を受けた雑誌によって認められること」「研究員ではない、バイアスのかかかっていない専門家による検討を受けること」がある。三者すべてにおいてデータ分析を用いることが求められ、一般性、再現性、汎用性が目的として掲げられているのである。しかし Reading First は、2.3.に示したとおり、無作為統制実験が含まれていない。さらに、「被験者の年齢、地理的、認知的、学力的、行動的特徴を示すこと」が求められ、その把握の上で「得られた知見はどういった人に対して一般化しうるのか示すこと」と、一般化の限界も前提として認識されている。この点において Reading First の実証性の基準と他の基準との間には大きな差異があるように考える。

そもそも、研究を実践に対置させたとき、その役割として期待されることは、研究成果の実践への応用、すなわち研究成果の一般性、汎用性、再現性である。しかしこれには限界があることも容易に推察されうる。なぜならば、研究成果が応用される現場の実践者（教師、子ども）は、一定の実践的傾向性を有していたとしても、自由意志を持つ存在である。したがって、ある介入が実践へ投入されたとしても、それに対して再現性が確保されるとおりに一連の行為（知覚、認識、実行等）が遂行されるとは限らない。一般性を求める以上、それには人間存在の自由意志が対置して存在しているのである。また、Reading First で「教育指導上の迫真性が確保されていること」と、実践と研究との間に一定の距離を置いている点は、整合性を有している。この人間存在の現実的側面を前提として考慮しないところに、意識的如何に関わらず、結果として恣意性を見出すことができる。

また、一般化に限界を見ていた Reading First が無作為統制実験を採用していない点も興味深い。無作為統制実験については先述のヴァレリ・レイナの解説にあるとおりであるが、この科学性を厳密に追究するためには、CSR に示されているように「主観的考えよりも事実即ちデータを用いること」「知覚可能なエビデンスを用いること」が求められることとなる。果たして人間の生きた実践を対象として、どこまで知覚可能な範囲で語りえ、主観性を排することが可能であろうか。本来ならば、主観性排除の前提の現実的可能性を探る必要性もある。

また、先に示した三者共通の項目として、「他者評価（ピアレビュー）を受けた雑誌によって認められること」「研究員ではない、バイアスのかかかっていない専門家による検討を受けること」があった。これは、真理性の観点からすると、専門的知見から客観的に再検討されることを意味すると同時に、本項の冒頭で述べた、語り主体が権威者であることによる、真理性の隠蔽の可能性に繋がる問題をも孕んでいる。ただ実態がそのどちらであるかは、検討の際用いられる実証性、真理性の基準といった論理構造によってしか確かめることはできない。

おわりに

本研究の究極的な目的は、教育行政領域における実証的研究について、実証性の基準（方法論）の正当性に対して、真理性の視点から改めて検討を加え、結果として、実証的研究が持ちうる権威性の正当性を検討することにある。そこで、本稿はその第一次的段階として、米国の NCLB 法を検討対象事例として取り上げ、そこで規定される「実証性」の基準を示すことを行った。その目的は、基準の正当性あるいは恣意性を、規定される基準の内から剔出することにあつた。そして、規定される基準に潜む恣意性として、人間存在の自由意志と一般性との両立困難性の問題を等閑視している点、ここで改めて本研究の重要性を提示しておきたい。そもそも、実証的研究が教育行政活動、教育実践活動上において必要であると認識されるようになった理由は何か。それは、学校における諸活動が、有効かどうかという点について実証的証拠を持たず、児童生徒の学力を改善させないプログラムやアプローチを実行してきたという連邦議会の疑念が伏在するからである。翻せば、公教育の質を改善する上では、教育行政活動、教育実践活動は実証的研究に基づくことが必要であり、また有効であること、良い研究は、よりよい教育を生むということ、こうした信念が、NCLB 法の中に実証的研究という言葉が 111 回も出現するという事実が意味するところなのである。

この信念が現実に妥当性、真理性をもつか否かを検討する必要性は十分にある。なぜならそれは、i) いかなる研究が公教育の質を改善させるのかといった、実証的研究に対する積極的課題においても、また ii) そうした政策が逆説的にもたらす余弊としての権威性の問題を検討するという消極的課題においても、重要だからである。本稿は ii) の課題意識の下に位置づけられるものだが、権威性の正当性に対して検討を行うためには、いまだ果たすべき課題が多く残存する。

1 つ目はすなわち、実証的研究が諸活動に対していかなる影響を及ぼしているか、どの程度一定の基準の枠の内に制約されているかについて、具体的事例を用いて考察することである。最も緩慢な基準しか要請しない Reading First は、最も権威性をもつことが制限されている。しかしこの最も緩慢な Reading First の基準においても、行政の解釈によっては十分に権威性を行使するに足る土台を持っている。実際、連邦が Reading First の実証的研究の基準を厳密に解釈しようとした結果、Reading First 補助金プログラムに対して州が出した申請の多くが、その実証性の問題から返却された事実がある。補助金を受給する地方学区は、実証性を基盤とする読解プログラムのみが資金を活用できること、そしてそれらプログラムのアプローチが確実に児童生徒の学力を向上させるというエビデンスが必要であると告げられたと言う²³。その権威性の妥当性については留保した上で、実際に諸活動がいかに制約されているのか、行政による解釈との関係性も踏まえ、知る必要がある。

2 つ目は、連邦政府自身の遂行能力の限界について検討することを意味する。上述のように、有効性、効率性の向上を意図して、教育行政活動、教育実践活動に実証的研究が組み込まれるのならば、常にその権威性は、正当な権威性として受容されるものでなければならない。それが不当な権威に墮するか否かは、実証的研究の基準を現実に遂行していく連邦政府自身の遂行能力に負うところである。実証的研究が本来の意図通り機能するためには、真理性に対峙し、権威者自身が、何が実証的であるか、真理であるかを判断できる能力を有していなければならない。連邦政府自身が真理性を前提として判断を遂行できなければ、実証的研究は、有効性、効率性という

点で、現実に機能しないものとなるのである。したがって、実証的研究の権威性については、1つ目の実践への影響の考察と並んで、連邦自身の遂行の実態及び遂行能力の限界を考察し、検討する必要がある。

そして3つ目は、冒頭で述べたように、基準の正当性、そしてその前提としてある真理性に対する哲学的考究である。本稿において示された、実証的研究であるための方法論、基準が、果たして現実に真理性を確保しているか否か、再検討する必要性を提示している。例えば、無作為統制実験や図表に示した基準に対して、一般的定義、CSRは盲目的に追従している感がある。実証性の基準が、不当な権威性に墮することを回避するためには、示される諸条件が、諸実践を制約するに十分に足る真理性を有しているのか、検討する必要がある。以上を今後の課題とする。

¹ 河野和清『現代アメリカ教育行政学の研究—行動科学的教育行政学の特質と課題』多賀出版、1995年、210頁。

² 河野和清、同上書、211頁。

³ 高木太郎「教育行政研究の手引き」山本敏夫・小川勝治・太田卓編『教育行政概説』御茶ノ水書房、1967年、318頁。

⁴ 中澤渉「教育社会学における実証研究の諸問題」『教育社会学研究』72集、2003年。

⁵ Pierre Bourdieu, *Choses Dites*, Editions de Minuit, 1987. (P. ブルデュール著、石崎晴己訳『構造と実践』藤原書店、1991年、14頁。)

⁶ 主として、Center on Education Policy, from the Capital to the Classroom, 2003, pp.101-116.

Jerry S. Carlson & Joel R. Levin ed., The No Child Left Behind Legislation: Educational Research and Federal Funding, A Volume in: Psychological Perspective on Contemporary Educational Issues, Information Age Publishing Inc., 2005 を用いた。

⁷ CSRとReading First以外にもImproving Teacher Qualityプログラムも、ガイダンスを提供し、その中で実証的研究について言及している。しかしそこで示されている実証的研究に関する定義は一般的定義と全く同じであり、特筆すべき点がないため、ここでは触れないこととした。

⁸ 2006年度は10億2923万4000ドル、2007年度は10億2923万4000ドルであった。

⁹ このプログラムは、もともと歳出予算の用語 (appropriations language) として設定され (Appropriations Act for the U.S. Department of Education, P.L. 105-78)、認可する法律 (authorizing legislation) ではなかった。

¹⁰ 2006年度は792万ドル、2007年度は235万1812ドルであった。

¹¹ ただし、2008年度から廃止され、他組織との連携も存在しない。

¹² プログラムによって、若干独自の定義をもつものがあるのは、連邦教育省内部では、各部署が互いに異なるプログラムを管理しており、したがって実証的研究の規定への対応の仕方も、プログラムによって異なってくるためである。

¹³ Coalition for Evidence-Based Policy, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Identifying and Implementing Educational Practices Supported By Rigorous Evidence: A User Friendly Guide, December 2003, p.1.

¹⁴ 2002年、Education Sciences Reform Act of 2002によって、OERIは廃止され、代わってInstitute of Education Sciences (IES) が連邦教育省内部に創設された。

¹⁵ Rockefeller Universityで、実験心理学を専攻し、博士号を獲得。当時のOERI局次長 (Assistant Secretary) Grover J. Whitehurst から、Senior Research Advisor に任命された。

¹⁶ Valerie Reyna, "What is Its Logic?", Department of Education, The use of Scientifically Based Research in Education, Working Group Conference, Wednesday February 6, 2002.

¹⁷ Coalition for Evidence-Based Policy, *ibid*, 2003, p.1.

¹⁸ Coalition for Evidence-Based Policy, *ibid*, 2003, p.2.

¹⁹ Coalition for Evidence-Based Policy, *ibid*, 2003, p.4.

²⁰ U.S. Department of Education, Office of Elementary and Secondary Education, Comprehensive School Reform (CSR) Program Guidance, 2002.

²¹ U.S. Department of Education, *ibid*, p.9.

²² U.S. Department of Education, Office of Elementary and Secondary Education, Guidance for the Reading First Program, 2002.

²³ Manzo K. K., "Majority of states told to revise reading plans" Education Week, October 2, 2002.

(比較教育政策学講座 博士後期課程 2 回生)

(受稿2009年9月7日、改稿2009年11月30日、受理2009年12月11日)

A Study on the Authority of Scientifically Based Research
Concerning Educational Practices:
The First Step Focusing on the No Child Left Behind Act of 2001

KIRIMURA Takafumi

The purpose of this paper is to examine the legitimacy of the standard of "scientifically based research". The No Child Left Behind (NCLB) Act seeks to improve the quality of public education by requiring educators to use programs and practices that are based on scientific research. The term "scientifically based research" appears 111 times in the NCLB Act. For all other programs under NCLB except for the Reading First program and the Comprehensive School Reform program, the law contains a highly specific definition of scientifically based research. The term "scientifically based research" includes research that is evaluated using experimental or quasi-experimental designs in which individuals, entities, programs, or activities are assigned to different conditions and with appropriate controls to evaluate the effects of the condition of interest, with a preference for random-assignment experiments. The U.S. Department of Education has issued guidance about implementing scientifically based research in the Reading First program and the Comprehensive School Reform (CSR) program. The definition of scientifically based research that applies only to the Reading First program as well as the definition of "strong evidence" in the CSR program is different from, and less rigorous than the definition that applies to all other programs in the NCLB Act.