

【論文】

米国連邦教育政策におけるエビデンス要求の現在 —全ての生徒が成功する法における4つのエビデンスの階層—

弘前大学 桐村 豪文

1. はじめに

近年、英国・米国をはじめ、また我が国も含め、教育において「エビデンスに基づく (Evidence-Based)」ことの是非が議論され、またその実施が要求されるようになった。例えば米国では、この種の要求が、学校改善を支援し奨励する連邦政府の取り組みに見られるようになったのは、1997年に成立した包括的学校改善デモンストレーションプログラムからであるが、その要求が最も顕著に現れたのは、2002年1月に成立したどの子ども置き去りにしない法 (No Child Left Behind Act: NCLB) である。その要求の出現の背景にあったのは、「連邦議会と連邦政府は、連邦補助金が、有効性についての幅広いエビデンスを欠く、児童生徒の学業成績を改善する見込みがほとんどない教育方略やアプローチを支援するために用いられていることに懸念を示し続けてきた」¹という事態であった。すなわち、限りある財源と競争の激しいグローバル社会を前に、国家の指示する方針に対して、教育研究、教育実践が期待に応える働きをしてこなかったという評価が、こうした動きの背景にはあるのである。

こうした要求の動きに対して、ありうる反応としては、この動きに同調して期待に応える働きをするか、あるいは、そもそもこの動きを疑問視し、これに抗い、まったく別の道を歩もうとするか、である。実際米国では、NCLB法によって「エビデンスに基づく」要求が大々的に出されたことを受けて、研究者の間では、研究の根底を支える主義主張をめぐる論戦が展開されたのである。中でも、前者の立場をめぐっては、それが科学主義、実証主義、あるいはポスト実証主義であるとして批判が起こり、その上で、それとの理論的差異を明示して自らの立場を再帰的に提示しようとする動きが様々あった²。

しかしながら、果たしてこの二分法もまた正しいのかという疑問を持つ必要がある。両端の間を歩む道はないのだろうか。つまり、国家の「エビデンスに基づく」政策意図を受け入れながら、と同時に、民主的で人間的な教育実践、教育研究を継続することができるという道である。本研究は、その遠大な道を探究するものである。

本稿では、その探究の一部として、米国における昨今の「エビデンスに基づく」政策の変化について考察を行う。前稿³では、NCLB法とCSRプログラム、i3を取り上げ、その変化を明らかにした(表1参照)。本稿では、その積み残しでもある、全ての生徒が成功する法 (Every Student Succeeds Act: ESSA) を扱う。同法では、後に詳述するが、それまでの厳格な科学的要請の方針を一変させ、ローカルな文脈を重視する方法がとられている。本稿では、その方針転換に、先の探究の道の一端を見ることができると考えている。なお、ESSAとNCLB法を比較する観点から、同法が全体的に「増大した連邦政府の権限拡大を制限、もしくは縮小を目指したもの」⁴であることはすでに北野が論じているが、「エビデンスに基づく」要求に特化してそうした比較考察を行うものは、管見の限り存在しない。

表1 「エビデンスに基づく」要求にかかわる連邦教育政策

年	出来事
1997	包括的学校改善デモンストレーション (Comprehensive School Reform Demonstration: CSRD) プログラムが成立。
2002	どの子ども置き去りにしない法 (No Child Left Behind Act of 2001: NCLB) が成立。 CSRSD は包括的学校改善 (Comprehensive School Reform: CSR) に名称変更。
2002	教育科学改革法 (Education Sciences Reform Act of 2002) が制定され、教育科学研究機構 (Institute of Education Sciences ; IES) が創設され、その中に教育エビデンス情報センター (What Works Clearinghouse: WWC) が設置される。
2009	米国再生・再投資法 (American Recovery and Reinvestment Act: ARRA) のもと Investing in Innovation Fund (i3) と Race to the Top (RTTT) が開始。
2015	全ての生徒が成功する法 (Every Student Succeeds Acts: ESSA) が成立。 Investing in Innovation Fund (i3) は Education Innovation and Research に名称変更。

【出典】筆者作成

以下、2節では、ESSAにおける「エビデンスに基づく」要求の具体的内容（「エビデンスに基づく」の定義、エビデンスの基準、連邦プログラムとの関係）について説明する。それを踏まえて3節では、旧法であるNCLB法との比較考察を行い、それを通してESSAの相対的特徴を明らかにしたいと思う。

2. ESSAにおける「エビデンスに基づく」要求

オバマ政権下の2015年12月10日に初等中等教育法 (Elementary and Secondary Education Act of 1965) の改正法として「全ての生徒が成功する法 (Every Student Succeeds Act: ESSA)」が成立した。ESSAでは、それまでの連邦政策と同様、エビデンスに基づく実践や方略、介入（以下、介入と総称する）の使用を強調している。またESSAでは、NCLB法の「科学的基盤をもつ研究 (scientifically based research)」の用語と同様、法律の条項を設けてその用語を定義している。それが表2である。

ESSAの第8101項(21)(A)では、エビデンスに基づく介入を、強力なエビデンス (strong evidence)、中程度のエビデンス (moderate evidence)、有望なエビデンス (promising evidence)、または理論的根拠を明らかにするエビデンス (evidence that demonstrates a rationale) によって裏付けられるものと定義している。すなわちESSAでは厳密さの水準に即して4段階のエビデンスを設けているのである。ただしこの定義だけでは各階層のエビデンスの内容が明確でないことから、連邦教育省は2016年9月に「州教育当局、地方教育当局、学校、教育者、および連携組織が個々の介入に関するエビデンスの厳密さを理解し、確認できるように」との趣旨のもと、「非規制的ガイダンス：エビデンスを用いて教育投資を強化する (Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments)」を発行している。これに従い、以下では、各階層のエビデンスの内容について詳しく見ていこうと思う。ただし、このガイダンスは拘束力をもたない「非規制的ガイダンス」であり、またガイダンスに「以下に示す基準は、ESSAに規定される4つの各水準にあるエビデンスを識別するため、連邦教育省が推奨する事項を示しています」と記されてあるように、あくまで「推奨事項」であることには留意を要する。

表2 ESSA に定められる「エビデンスに基づく」の定義

<p>Section 8101 (21) エビデンスに基づく (EVIDENCE-BASED)</p> <p>(A) 一般的事項</p> <p>副段落 (B) に記載されている場合を除き、「エビデンスに基づく」という用語は、州教育当局、地方教育当局または学校の活動に関して使用される場合、以下の特徴を有する活動、方略または介入を意味する。</p> <p>(i) 児童生徒に係る結果またはその他の関連する結果の改善に対して、以下に示すエビデンスに基づき、統計的に有意な効果を示していること。</p> <p>(I) 少なくとも1つの、適切に設計され、適切に実施された実験的研究から得られる強力なエビデンス。</p> <p>(II) 少なくとも1つの、適切に設計され、適切に実施された準実験的研究から得られる中程度のエビデンス。または</p> <p>(III) 選択バイアスを統計的にコントロールしたもので、少なくとも1つの、適切に設計され、適切に実施された相関研究から得られる有望なエビデンス。または</p> <p>(ii) (I) 当該活動、方略または介入が、児童生徒に係る結果または他の関連する結果を改善する可能性が高いことを示す、質の高い研究結果または肯定的評価に基づく、理論的根拠を明らかにするもの。そして</p> <p>(II) 当該活動、方略または介入の効果を調べるための継続的な取り組みが含まれる。</p> <p>(B) この法律の下で補助を受ける特定の活動に対する定義</p> <p>第1003項に基づき補助を受ける介入、改善活動または方略に関しては、「エビデンスに基づく」という用語は、州教育当局、地方教育当局または学校の活動、方略または介入のうち、副段落 (A) (i) の副条項 (I) (II) または (III) の要件を満たすものを意味する。</p>

【出典】 Every Student Succeeds Act, Public Law 95, 114th Cong. 1st sess. (December 10, 2015), pp. 289–290, Retrieved from <https://www.congress.gov/bill/114th-congress/senate-bill/1177/text>.

2. 1. ESSA における「エビデンス」が満たすべき基準

(1) 強力なエビデンス (Strong Evidence)

強力なエビデンスに裏付けられるには、介入に関し、適切に設計され、適切に実施された実験研究(例、ランダム化比較試験 (randomized control trial)) が少なくとも1つ必要とされる。連邦教育省は、留保なしの WWC エビデンス基準 (WWC Evidence Standards without reservations) を満たしているか、因果推論を行う点でそれと同等の質をもったものである場合、実験研究を「適切に設計され、適切に実施されている」と見なす。さらに、強力なエビデンスをもたらすために、その研究では以下のことを行う必要があるという。

- 1) 児童生徒に係る結果またはその他の関連する結果 (relevant outcome) に関して、介入の統計的に有意で肯定的な (つまり、好ましい) 効果を示していること。
- 2) 留保ありまたは留保なしの WWC エビデンス基準 (WWC Evidence Standards with or without reservations) を満たす、または因果推論を行う点で同等の質をもったものである他の研究におけ

る、同介入に関する統計的有意で否定的な（すなわち、好ましくない）エビデンスによって無効にされないこと。

- 3) 大規模なサンプル（large sample）で複数の場所で開催されたサンプル（multi-site sample）があること。
- 4) 介入を受けるよう提案がなされた母集団（すなわち、介入の対象とする児童生徒のタイプ）および環境（農村部、都市部など）において重複のあるサンプルを用意すること。

（2）中程度のエビデンス（Moderate Evidence）

中程度のエビデンスに裏付けられるためには、介入に関し、適切に設計され、適切に実施された準実験的研究（quasi-experimental study）が少なくとも1つ必要とされる。連邦教育省は、留保ありの WWC エビデンス基準（WWC Evidence Standards with reservations）を満たしているか、因果推論を行う点でそれと同等の質をもったものである場合、準実験的研究を「適切に設計され、適切に実施されている」と見なす。さらに、中程度のエビデンスをもたらすために、その研究では以下のことを行う必要があるという。

- 1) 児童生徒に係る結果またはその他の関連する結果に関して、介入の統計的有意で肯定的な（つまり、好ましい）効果を示していること。
- 2) 留保ありまたは留保なしの WWC エビデンス基準を満たす、または因果推論を行う点で同等の質をもったものである研究における他の研究結果から得られる、その介入に関する統計的有意で否定的な（すなわち、好ましくない）エビデンスによって無効にされないこと。
- 3) 大規模なサンプルで複数の場所で開催されたサンプルがあること。
- 4) 介入を受けるよう提案がなされた母集団（すなわち、介入の対象とする児童生徒のタイプ）または環境（農村部、都市部など）において重複のあるサンプルを用意すること。

（3）有望なエビデンス（Promising Evidence）

有望なエビデンスに裏付けられるためには、介入に関し、選択バイアスに対する統計的コントロールを備えた、適切に設計され適切に実施された相関研究が少なくとも1つ必要とされる。連邦教育省は、介入群と比較群との間の相違を低減させるまたは考慮するようなサンプリングおよび／または分析方法を使用する場合、相関研究を「適切に設計され、適切に実施された」とみなす。さらに、有望な証拠をもたらすために、その研究では以下のことを行う必要があるという。

- 1) 児童生徒に係る結果またはその他の関連する結果に関して、介入の統計的有意で肯定的な（つまり、好ましい）効果を示していること。
- 2) 留保ありまたは留保なしの WWC エビデンス基準を満たす、または因果推論を行う点で同等の質をもったものである研究における研究結果から得られる、その介入に関する統計的有意で否定的な（すなわち、好ましくない）エビデンスによって無効にされないこと。

（4）理論的根拠を明らかにする（Demonstrates a Rationale）

理論的根拠を明らかにするために、介入には以下のものを含める必要がある。

- 1) どうして介入が関連する結果を改善する可能性が高いかを示す研究または評価からの情報をもとに、十分明確に記したロジックモデル (logic model)。
- 2) 理想的には、有望なエビデンスまたはそれ以上のエビデンスを生むもので、その介入の成功について利害関係者に情報を与えるために、介入の一部として行われる、または他の場所で進行中の (例えばこれは、別の州教育当局、地方教育当局または研究機関が他の場所で介入を研究していることを意味する可能性がある)、介入の効果を研究する取り組み。

2. 2. ESSA における 4 つの階層のエビデンスの適用

以上、ESSA に定める 4 つの階層のエビデンスについて、それぞれが満たすべき基準の内容について見てきた。表 3 は、それらの要点を一覧表にまとめたものである。

なお、上の説明にあった「留保なしの WWC エビデンス基準 (WWC Evidence Standards without reservations)」と「留保ありの WWC エビデンス基準 (WWC Evidence Standards with reservations)」は、2002 年に連邦教育省に教育科学研究機構 (Institute of Education Sciences) が創設され、その中に設置された教育エビデンス情報センター (What Works Clearinghouse: WWC) によるエビデンスの評価に関わるものである。WWC は、2008 年にエビデンス基準 (Evidence Standards) を策定し、これまで 3 度改訂している (現在、Standards Handbook, Version 4.1 が発行されている)。

そのうえで、前者の「留保なしの WWC エビデンス基準」は、連邦教育省の説明によれば、WWC によってレビューされたグループデザイン研究の最高の評価であり、この評価を受けた研究は、観察された効果が介入によって引き起こされたという最高の信頼度を提供する。適切に実施されたランダム化比較試験 (すなわち、脱落 (attrition) のような問題によって損なわれないランダム化比較試験) は、この最高の評価を受ける可能性がある、という。そしてもう一つの「留保ありの WWC エビデンス基準」は、WWC がレビューしたグループデザイン研究の中程度の評価であり、この評価を受けた研究は、観察された効果が介入によって引き起こされたという信頼度が低くなる。適切に実施されていないか、脱落の問題があるランダム化比較試験は、強力な準実験的デザインとともに、この評価を受ける可能性がある、という。

ESSA における「エビデンスに基づく」要求の最大の特徴は、上に示したとおり、エビデンスに 4 つの階層が設けられたことであり、留保なしの WWC エビデンス基準を満たす階層 1 の強力なエビデンスから、ロジックモデルを含む階層 4 の理論的根拠を明らかにするものまで、幅広く「エビデンス」として認めていることがわかる。

ただし、その適用をめぐるのは、次のような規定となっている。

初等中等教育法のタイトル I、II、IV のもと、連邦補助を受ける学校は、「エビデンスに基づく」介入を実施することが規定されている。その中でもタイトル I のもと、包括的支援及び改善 (comprehensive support and improvement: CSI)、焦点化した支援及び改善 (targeted support and improvement: TSI)、追加的な焦点化した支援及び改善 (additional targeted support and improvement: ATSI) を要すると特定された低成績学校については、階層 1 の強力なエビデンス、階層 2 の中程度のエビデンス、階層 3 の有望なエビデンスのいずれか (階層 4 は認められない) を使用することが義務付けられている。それ以外の場合は、4 つの階層のいずれかのエビデンスを示す必要がある (表 4 参照)。

表3 各階層のエビデンスが満たすよう推奨される基準の概要

	強力なエビデンス (Strong Evidence)	中程度のエビデンス (Moderate Evidence)	有望なエビデンス (Promising Evidence)	理論的根拠を明らかにする (Demonstrates a Rationale)
研究デザイン	実験的研究	準実験的研究	選択バイアスに対する統計的コントロールを備えた相関研究	研究または評価からの情報をもとに、十分明確に記したロジックモデルを提供するもの
WWC基準	留保なしのWWCエビデンス基準を満たす（または、同等の質をもつ）	留保ありまたは留保なしのWWCエビデンス基準を満たす（または、同等の質をもつ）	該当なし	該当なし
好ましい効果	児童生徒に係る結果またはその他の関連する結果に関して、介入の統計的有意で肯定的な（つまり、好ましい）効果を示していること	児童生徒に係る結果またはその他の関連する結果に関して、介入の統計的有意で肯定的な（つまり、好ましい）効果を示していること	児童生徒に係る結果またはその他の関連する結果に関して、介入の統計的有意で肯定的な（つまり、好ましい）効果を示していること	関連する研究または評価が、介入が児童生徒に係る結果またはその他の関連する結果を改善する可能性が高いかを示していること
他の効果	留保ありまたは留保なしのWWCエビデンス基準を満たす（または、同等の質をもつ）研究における他の研究結果から得られる、統計的有意で否定的な（すなわち、好ましくない）エビデンスによって無効にされないこと	留保ありまたは留保なしのWWCエビデンス基準を満たす（または、同等の質をもつ）研究における他の研究結果から得られる、統計的有意で否定的な（すなわち、好ましくない）エビデンスによって無効にされないこと	留保ありまたは留保なしのWWCエビデンス基準を満たす（または、同等の質をもつ）研究における他の研究結果から得られる、統計的有意で否定的な（すなわち、好ましくない）エビデンスによって無効にされないこと	理想的には、有望なエビデンスまたはそれ以上のエビデンスを生むもので、介入の一部として行われる、または他の場所で進行中の、介入の効果を研究する取り組み
サンプルサイズと重複	介入を受けるよう提案がなされた母集団および環境において重複のある、大規模なサンプルと複数の場所で実施されたサンプルを含むこと	介入を受けるよう提案がなされた母集団または環境において重複のある、大規模なサンプルと複数の場所で実施されたサンプルを含むこと	該当なし	該当なし

【出典】 U.S. Department of Education, ESSA Guidance using evidence to strengthen education investments, 2016, Retrieved from <https://www2.ed.gov/policy/elsec/leg/essa/guidanceusesinvestment.pdf>.

表4 ESSA の各プログラムに対応するエビデンス要件の概要

連邦プログラム	要求されるエビデンス
Title I, Section 1003: School Improvement	CSI、TSI、および ATSI の各学校で少なくとも1つの介入が、階層1、2または3のエビデンス要件を満たす必要がある。
Title I, Part A: Basic Programs for Schoolwide and Targeted Assistance	学校がスクールワイドプログラムまたは焦点化支援プログラムの実施を支援してくれるよう選定した外部の供給業者は、エビデンスに基づく方略の使用に関する専門知識を持たなければならないが、それ以外ではエビデンスの最低基準値といったものはない。
Title II, Part A: Supporting Effective Instruction and Literacy for All	補助金の特定の使用（専門職開発、任命、メンタリングなど）については、当該エビデンスが合理的に入手可能と州が判断する限りにおいて、階層1、2、3または4のエビデンスが必要である。識字率向上を目的とする場合は、州は、階層1、2または3を満たす団体に対して補助を優先的に与えなければならない。
Title IV, Part A: Student Support and Academic Enrichment Grants	補助金の特定の使用には、州が、当該エビデンスが合理的に入手可能であると判断する限りにおいて、階層1、2、3または4のエビデンスが必要である。
Title IV, Part B: 21st Century Community Learning Centers	適切と見なされる場合、階層1、2、3または4のエビデンスが必要である。
Title IV, Part D: Magnet School Assistance	エビデンスに基づく活動（階層1、2、3または4）を伴う提案には、競争上優先権が与えられる。
Title IV, Part F: Education Innovation and Research	提案されたイノベーションは、階層1、2、3または4のエビデンスを満たす必要がある。
Title IV, Part F: Community Support for School Success	<p>Promise Neighborhoods 申請要件には、節目（例：幼稚園に入る前～第12学年から中等教育後）を経験する児童生徒を援助する EBP の支援が含まれており、エビデンスに基づく活動（階層1、2または3）の提案に対しては競争上優先権が与えられる。</p> <p>Full-Service Community Schools エビデンスに基づく活動（階層1、2または3）を伴う提案には、競争上優先権が与えられる。</p>

【出典】 Alicia Garcia, Elisabeth Davis, ESSA Action Guide: *Selecting Evidence-Based Practices for Low-Performing Schools*. American Institutes for Research, 2019, Retrieved from <https://www.air.org/resource/essa-action-guide-selecting-evidence-based-practices-low-performing-schools>.

なお、タイトルⅡでは、一部のプログラムがエビデンスに基づく必要があり、また一部のプログラムはすべての要素ではなく一部の要素についてエビデンスに基づくことが規定され、また一部のプログラムはエビデンスに基づく規定はあるが、エビデンスが十分に存在しない場合は、州によって免除されることもある⁶。

3. ESSAにおける「エビデンスに基づく」要求の特徴

ESSAにおける「エビデンスに基づく」要求の特徴を、従来の政策（NCLB法）と比較しながら、またとりわけ懸念される事項に焦点を当てながら、明らかにしたい。

3. 1. NCLB法との比較

エビデンスに基づく学校改善を支援し奨励する連邦政府の取り組みは、冒頭に示したとおり、NCLB法に遡ることができる。NCLB法では、州や学区に対し、「科学的基盤をもつ研究（scientifically based research）」によって裏付けられた介入を使用することが求められた。そこで認められるエビデンスは、原則として、適切に実施されるランダム化比較試験または準実験デザインを用いた質の高い研究（ESSAにおける階層1と階層2のエビデンスにほぼ一致するもの）に限られていた。

そもそもNCLB法には4本柱が打ち立てられてた。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）がそれである。しかしながら、③に限ってその当時のインパクトを見ると、教育政策センター（Center on Education Policy）が2002年の夏から秋にかけて州教育当局者に行ったインタビューによれば、「私たちのインタビューでは、ほとんどの州のタイトルⅠ監督者は、科学的基盤をもつ研究の要件とその潜在的な重要性は理解していたものの、それを履行するためのことはほとんど何もしなかったことが明らかとなった。これまでのところ、州のタイトルⅠを監督する部局は、アセスメントやアカウントビリティといったすぐにでもアクションを要するNCLB法の他の規定により焦点を当てており、科学的基盤をもつ研究の要件にほとんど注意を払っていない。」⁷という実態が示されたのである。そしてまた、「私たちがインタビューした多くの州は、科学的基盤をもつ研究についての法律の定義を厳密に適用すると、それを完全に満たすプログラムやアプローチを見つけることが難しくなることを懸念していた。例えば、定義は『ランダム化』の研究を好んでおり、その研究では、大勢の児童生徒が2つのグループにランダムに割り当てられる。1つのグループでは、検証中のプログラムや教育方法を受け、もう1つのグループでは、学校が通常使用するプログラムや方法を受ける。これらの類の研究は、費用がかさみ、学校の状況の中にこうした条件を設定するには複雑であり、家族が子どもの参加に同意するのが難しいため、教育ではあまり一般的ではない。また、子どもは研究の対象にはない別のところから教育サービスを受けている可能性があり、児童生徒の流動性は処置群と統制群の両方に影響するため、統制群を維持することも困難である。さらに、子どもにとって有益である可能性のあるサービスを認めない、ということ正当化するのは無理がある。」⁸という声が州教育当局者から上がっており、NCLB法が求める「エビデンス」の定義の狭さとそのローカルな文脈から乖離した現実性の薄さが2002年当時から指摘されていたのである。

表5 NCLB法の条項を連邦教育省がどの程度厳格に実施しているかについての評価と州の数

	非常に厳しい	厳しい	やや厳しい	まったく厳しくない
Adequate yearly progress	25	16	7	1
Reading First	22	18	2	0
Public school choice	14	14	11	4
Supplemental education services	11	19	11	2
Highly qualified teachers	8	20	16	2
Paraprofessional qualifications	8	12	21	1
Scientifically based research	4	11	17	9

【出典】 Diane Stark Rentner, et al., From the Capital to the Classroom: Year 3 of the No Child Left Behind Act, Center on Education Policy, 2005, p.73.

また、2004年12月に教育政策センターが州に対して行ったアンケート調査（49州が回答）によれば、表5に示すとおり、NCLB法に定める諸条項の中で、連邦教育省が厳格に実施していると州が感じているものとして、Adequate yearly progress や Reading First が挙げられるが、科学的基盤をもつ研究（Scientifically based research）についてはこの中で最下位に位置し、州にとって、あるいは（州から見た）連邦政府にとって、相対的に重視されるものではなかったということが伺える⁹。

結局のところ、NCLB法で要求するこの極めて狭く厳しい基準により、それを満たす研究が不足したため、州や学区がエビデンスに基づく学校改善に真に取り組むことは困難になり、そのため、この種の取り組みはほとんど成果を生むことはなかった¹⁰。

対照的に、ESSAでは、低成績学校を改善するために許容可能なエビデンスの範囲を広げるという、従来と異なるアプローチをとる。すなわち、階層1から階層4までの幅をもって、目下の教育研究基盤の現実と意思決定者が利用可能な選択肢により適する形で、「エビデンスに基づく」政策を進めようとするものである。

また、現実の実態に即してエビデンスの幅を広げたということに加え、こうしてエビデンスの範囲を広げたことにより、州や学区の当局者は、連邦政府の補助を受けて実施するアプローチをより柔軟に検討できるようになったのである。階層3や階層4（タイトルIについては階層4は認められないが）のエビデンス要件が用意されることにより幅広い選択肢が可能となり、それは、ローカルなニーズや改善計画に最適なアプローチを見つけるのに役立つのである¹¹。なお、ESSAでは前述の通り、階層化されたエビデンスの基準が示されているが、しかしその階層性は、いずれかを優先させようとする政策意図を含んだものではない。NCLB法では、エビデンスの基準にはっきりとした階層性が存在し、その上に、政策上、認められるものと認められないものとの境界が明確化されていた。一方、ESSAでは、階層性は提示されながらも、政策運営上、認められるべきエビデンスの基準に厳しい限定は加えられていない。連邦教育省によるガイダンスでは、「州教育当局、地方教育当局、およびその他の利害関係者は、(略)可能であれば、より高いレベルのエビデンス、特に因果推論を通して介入の有効性を説明する、強力なエビデンスや中程度のエビデンスによって裏付けられる介入を優先させる必要があります」¹²と述べら

れているが、あくまでそれは推奨事項であり、したがって例えばタイトルIにおいても、階層1の「強力なエビデンス」から階層3の「見込みのあるエビデンス」まで、それ以外のプログラムでは、階層1の「強力なエビデンス」から階層4の「理論的根拠を明らかにするもの」まで、どのエビデンスの基準を満たす介入を使用しようと認められるのである。

なお、ESSAの可決以降、投稿された多くのブログや解説、記事では、階層4のエビデンスの定義の曖昧さに対して懸念が強調されている¹³。次項では、その階層4のエビデンス「理論的根拠を明らかにするもの」について、Rebecca Herman et al. (2016)をもとに、より詳しく見ていこうと思う。

3. 2. 階層4のエビデンス「理論的根拠を明らかにするもの」とは

Rebecca Herman et al. (2016)によれば、連邦教育省や州教育当局、地方教育当局が階層4のエビデンスの定義をどのように解釈するかによって、以下に示す研究がエビデンスと見なされたり見なされなかったりする、という。

例えば、構成要素はエビデンスに基づいているが、介入全体はその有効性が証明されていないもの。これには、介入が全体として有効であることを証明する質の高いエビデンスがない場合に、介入の構成要素（例、教室で発揮される専門能力の開発）が児童生徒に係る結果またはその他の関連する結果を改善することを発見する質の高い研究が含まれる。介入の開発者は、効果的な構成要素を使用しているから自身の介入は効果的なのだと主張するものの、これは、WWCが創業して5年の間、繰り返し区別していたことである。介入が新しいか、またはコンテキストに合わせて改良させている場合、完全な状態の介入の効果についてのエビデンスを得ることはできないかもしれない。構成要素が効果的であることを知ることは、介入が全体として効果的であるかもしれないという指標を提供する。しかし、これは、実証された効果を示す強力なエビデンスからは程遠いということ、また認識するのである。

また、因果連鎖の有効性についての研究。これには、中間に介在する結果（例、教師の内容知）が児童生徒に係る結果またはその他の関連する結果に影響するというを示した研究結果の上に、介入が中間に介在する結果（例、教師の内容知）に影響を与えるというを示す質の高い研究が含まれる。

既存のエビデンスと実践におけるニーズとの間に間隙がある中で、以上に挙げた研究はすべて、WWCのエビデンス基準は満たさないにせよ、実践に対して貴重な情報を提供するものである。例えば、ある介入の中心的構成要素はそれ単体では効果を示すエビデンスがあり、そして他の構成要素と組み合わせると介入を実施することができるというエビデンスが示されたならば、まだその有効性が実証されていない介入に対して、強力的に弁護するものとなるのである¹⁴。

Rebecca Herman et al. (2016)によれば、以上に挙げた研究のいずれも、階層4のエビデンスとして認められる可能性があるという。ただし、階層4としても認められないものと見なされるエビデンスには、次のようなものが含まれるという。すなわち、改善活動の成功に関して故意に選んだ逸話、検証を受けておらず、的外れであり、あるいはその正当性が立証されていない結果の分析（例、世論調査）、結果の分析なしに提示された理論、である。

ESSAでは、先述の通り、現実の実態に即してエビデンスの幅を広げたものの、「何でもあり」というわけではないようである。だが、とりわけ階層4のエビデンスについては曖昧性がなお残っており、その点は、「幅広い解釈に開かれている」¹⁵といえる。

4. おわりに

本稿では、米国における昨今の「エビデンスに基づく」政策の変化について考察を行った。2015年12月10日に成立した全ての生徒が成功する法（ESSA）では、エビデンスに、強力なエビデンス、中程度のエビデンス、有望なエビデンス、理論的根拠を明らかにするエビデンスの4つの階層を設けている。ESSAの特徴は、このエビデンスの幅の広さである。旧法のNCLB法では、原則としてランダム化比較試験または準実験法を用いた研究のみがエビデンスとして認められていたのだが、それと対照的である。そして、このエビデンスの幅の広さは、州や学区の当局者が、連邦政府の補助を受けて実施するアプローチを検討する際、幅広い選択肢が可能となり、極めて狭く厳しい基準に縛られることなく、ローカルなニーズや改善計画に適したアプローチを見つけることの援助の一つとなるのである。

冒頭に述べた通り、本稿は、国家の「エビデンスに基づく」政策要求と、民主的で人間的な教育実践、教育研究の継続の両者を同時に満たす可能性の探究を目的として、その手がかりを米国連邦政策における昨今の方針転換に求め、その考察を行ったものである。本稿で明らかとなったのは、NCLB法の政策の失敗と、その背景にあった「エビデンス」の定義の狭さとそのローカルな文脈から乖離した現実性の薄さであった。ESSAはまさにその反省の上に、「エビデンスに基づく」政策の要求においては、同時にローカルな文脈を考慮できるよう、制度設計（階層1から4のエビデンスの幅の広さと、幅広い解釈に開かれた階層4のエビデンス）が為されているのである。子ども一人ひとりの学びとそれを支える教育が、部分最適に安住することなく全体最適に向けて進み続ける動力を持ち続けるためには、国家の役割の重要性はもちろんのこと、しかしそこでは同時にローカルな文脈に込められる仕組みを用意することが重要なのではないだろうか。ただしそれは単に、国家の役割のいくつかを削減する（政府権限の制限、縮小）ものと捉えるのでは不十分である。国家の「エビデンスに基づく」政策要求と、民主的で人間的な教育実践、教育研究の継続の両者を同時に満たす可能性を見出すうえでは、国家の役割はより戦略的でなければならない。本稿で提示できたことは、こうした構想の一縷の可能性である。

とはいえ、本稿はなおもいくつもの研究課題を残している。米国国家の「エビデンスに基づく」政策の方針転換の背景には、果たしてどのような意図があり、具体的にどのような考えや実態があったのか。そしてまた、このような方針転換によって、果たして本当に「ローカルなニーズや改善計画に適したアプローチを見つけること」ができていたのだろうか。また、このような「エビデンスに基づく」要求の柔軟化は、「児童生徒の学業成績を改善する見込みのある教育方略やアプローチを支援する」という趣旨に近づくものなのだろうか、あるいはむしろそこから遠ざかるものなのだろうか。もしこの柔軟化という方針だけでその趣旨に近づくことはないとするならば、ではこれに加えて他にどのような手立てを要するのだろうか。こうした問いについては、また別稿で論じることにはしたい。

¹ Diane Stark Rentner, Naomi Chudowsky, Tom Fagan, Keith Gayler, Madlene Hamilton, Nancy Kober, Elizabeth Pinkerton, Jack Jennings, Caitlin Scott, Barbara Buell, *From the Capital to the Classroom: State and Federal Efforts to Implement the No Child Left Behind Act*, Center on Education Policy, 2003, p.104.

² 桐村豪文「因果関係の発見をめぐる認識論的・方法論的立場の多様性—米国における「科学的な教育研究」の内容をめぐる議論を参考にして—」『教育行財政論叢』12号、2013年、69-86頁。

-
- ³ 桐村豪文「米国連邦教育政策におけるエビデンス要求の変遷—エビデンス要求の厳格化から緩和に向けて」『地域連携教育研究』3号、2018年、35-47頁。
- ⁴ 北野秋男「オバマ政権の教育改革—RTTT政策からESSA法まで—」『国際教育』23巻、2017年、1-16頁。
- ⁵ U.S. Department of Education, *ESSA Guidance using evidence to strengthen education investments*, 2016, p.7.
- ⁶ Rebecca Herman, Susan M. Gates, Emilio R. Chavez-Herrerias, Mark Harris, *School Leadership Interventions Under the Every Student Succeeds Act: Volume I—A Review of the Evidence Base, Initial Findings*, The RAND Corporation, 2016, p.5. Retrieved from https://www.rand.org/pubs/research_reports/RR1550.html.
- ⁷ Diane Stark Rentner, et al., op. cit., 2003, p.110.
- ⁸ Ibid., p.114.
- ⁹ Diane Stark Rentner, Nancy Kober, Caitlin Scott, Naomi Chudowsky, Victor Chudowsky, Scott Joftus, Linda Carstens, *From the Capital to the Classroom: Year 3 of the No Child Left Behind Act*, Center on Education Policy, 2005, p.73.
- ¹⁰ Martin R. West, “From evidence-based programs to an evidence-based system: Opportunities under the Every Student Succeeds Act”, 2016, Retrieved from www.brookings.edu/research/from-evidence-based-programs-to-an-evidence-based-system-opportunities-under-the-every-student-succeeds-act.
- ¹¹ Steve Fleischman, Caitlin Scott, and Scott Sargrad, *Better Evidence, Better Choices, Better Schools. State Supports for Evidence-Based School Improvement and the Every Student Succeeds Act*, Center for American Progress, 2016, p.6.
- ¹² U.S. Department of Education, op. cit., 2016, p.8.
- ¹³ Advanced Education Measurement, “Research & the ESSA Pt 2: The Four-Tiered Problem”, 2016, Retrieved from <http://www.advancededucationmeasurement.com/commentary/research-the-essa-pt-2-the-four-tiered-problem>; Robert E. Slavin, “Evidence and the ESSA,” Huffington Post, 2015, Retrieved from http://www.huffingtonpost.com/robert-e-slavin/evidence-and-the-essa_b_8750480.html; Martin R. West, op. cit., 2016.
- ¹⁴ Rebecca Herman, et al., op. cit., 2016, pp.7-8.
- ¹⁵ Ibid., p.6.

Current State of "Evidence-Based" Policy in the United States in Education: The Four Tiers of Evidence in the Every Student Succeeds Act

Takafumi KIRIMURA

This paper examines the recent changes in the “evidence-based” policy in the United States, and explores the possibility whether the government’s intention of the “evidence-based” policy and the continuance of the democratic and humanized teaching practices and research could take place at the same time.

The Every Student Succeeds Act (ESSA), enacted on December 10, 2015, emphasized the use of evidence-based practices, strategies and interventions, as in previous federal policies. The Act defines evidence-based intervention as supported by strong evidence, moderate evidence, promising evidence, or evidence that demonstrates a rationale. In other words, ESSA provides four tiers of evidence according to the level of rigor.

The primary feature of ESSA is the relative breadth of the legally accepted evidence. The ESSA stands in contrast to the No Child Left Behind Act (NCLB), where for NCLB, only studies using randomized controlled trials or quasi-experimental methods were accepted as evidence. And the breadth of the ESSA’s evidence means that state and school district officials have a wide range of options when considering federally funded approaches, instead of being bound by extremely narrow and strict standards. It is one of the aids in finding the appropriate approach for local needs and improvement plans.

However, concerns about the definition of Tier 4 evidence have been emphasized from a number of perspectives as ambiguous, and in fact its definition is open to widespread interpretation.