

アメリカにおける科学的探究力の評価に関する検討

——『2009年のNAEPのための科学の枠組み』に焦点を合わせて——

大 貫 守*

1. はじめに

知識基盤型社会において生きる力や21世紀型コンピテンシーなどの教科汎用型のスキルの育成が重視されている。日本の理科教育では、単なる科学的概念などの内容の習得を超えて、科学的探究力といった能力を育成することが企図されている。この点について、例えば、2018年改訂の高等学校学習指導要領では、理数探究などの探究的な科目が新設されている。加えて、大学入試改革に向けた議論では、科学的探究力を評価する方法の開発が喫緊の課題となっている。

評価研究の文脈においても、このような科学的探究力を評価するためのオルタナティブな評価方法が提案されている。例えば、二宮衆一（2019）は、京都市立堀川高等学校の探究的な学習の事例などを挙げながら、子どもたちの行為の質を判断するための評価基準表であるルーブリックを用いた評価の取り組みを紹介している。また西岡加名恵（2016）は、先の京都市立堀川高等学校と同様に研究開発指定を受けて、理数教育に特化した教育を行うスーパー・サイエンス・ハイスクール（Super Science Highschool：以下、SSH）などにおけるポートフォリオとルーブリックを併用した探究力の評価の可能性に言及している。

実践レベルで、このような科学的探究力に関する評価の研究を担ってきたのが、上述のSSHである。多くのSSHでは科学的に探究する力などを身につけることが教育目標として掲げられ、課題探究型の授業が実施されている。その上で、その授業を通して獲得した科学的探究力を評価する方法についても研究が進められている。具体的には、子どもたちの成長の道筋を明らかにするルーブリックや自己評価によるアンケート

等の開発が挙げられる（例えば、SSH連絡会、2018）。

以上から日本の科学的探究の評価研究の動向について、次の2つの特徴を指摘できる。まず、国内の科学的探究力の評価に関する研究は、主に学校や教室の文脈で構想されている。そこでは、教師の指導や子どもたちの学習の改善が意図され、より個別具体的文脈に根ざした評価方法の開発や普及が行われている。

それに付随して、2点目として多くの学校で真正の評価が行われ、ルーブリックを用いて取り組みを質的に判断するという方法が使われている。その中では、独自の観点として教員自身の経験に裏打ちされた理論に立脚し、課題設定力などが位置づけられている。

このような教室や学校レベルでの評価は、個人や学校の置かれている状況に応じて、子どもたちの長期的な成長を教師が支え励ますものとして機能する一方で、次の課題も指摘できる。まず、①評価が実践現場で培われた鑑識眼に過度に依拠し、何をどのように評価するのかという点について、その理論的な裏付けが不明瞭なものとなることがある。次に、②社会人基礎力など、一見すると科学的探究力とは直接的には関わりのない力がルーブリック等の評価規準（基準）の根拠として参照される事例もみられる（大久保・森・中切、2018）。その結果、局所的で恣意的な評価となり、評価の比較可能性を担保できないことや科学的探究力の評価として、構成概念妥当性が確保されない危険性がある。

この点について、科学的探究を科学教育の主たるテーマとして掲げてきたアメリカでは、科学的探究力に関して、何をどのように評価してきたのだろうか。本稿では、全米規模で比較可能な形で行われている大

規模調査である全米学力調査 (National Assessment of Educational Progress: 以下、NAEP) に着目してこの点を検討する。具体的には、2009年に開発された科学の問題開発の枠組みである『2009年のNAEPのための科学の枠組み (Science Framework for the 2009 National Assessment of Educational Progress: 以下、Framework)』に即して、科学的探究の評価について客観的な指標として何が選び取られ、どのような手法で評価されてきたのかということの詳細にする。

NAEPの全体の概要については、荒井克弘ら (2008) によって既に日本に紹介されている。同研究では、NAEPの枠組みやその設立の背景、2000年代の調査結果についてまとめている。その中でも、安野史子 (2008) が、2000年代前半の科学の評価の枠組みに焦点を合わせて概説している。しかしながら、これは枠組み全体の構成を明らかにすることに主眼が置かれており、その評価の枠組みについて科学教育の議論に即して、その是非を問うものではない。

他方で、NAEPのFrameworkについては、アメリカの科学教育の研究者らが分析している。例えば、丹沢哲郎 (2012) は2008年に全面改訂されたNAEPのFrameworkについて、改訂の背景、評価の対象となる内容領域、その習熟のレベルについて概説している。また、徳永聖一・坂本憲明 (2014) では、2009年に実施された実際の問題例と採点基準を挙げている。

特に、NAEPの科学的探究力の評価に関する研究に、石崎友規 (2013) がある。石崎は、第4学年の問題解決能力を評価するための課題として、実際に実験活動を行わせるハンズ・オンの課題 (hands-on tasks) やコンピュータを用いたシミュレーション課題の内容について具体的に示している。その上で、これらの課題は子どもを科学の問題解決の場面に立たせ、そのパフォーマンスを評価するものであると言及している。

一連の研究は、確かにNAEPの評価の具体を詳らかにするものである。しかしながら、科学的探究の評価に関する分析という点では必ずしも十分ではない。この点についてアメリカの科学的探究力の評価の研究では、科学的探究力の評価のアプローチの多様性は「科学的探究の概念の定義と用いられている [科学的探究力を] 引き出す方法の種類」に由来するものであると既に指摘されている (Hanauer, Hatfull, & Jacobs-sera, 2009, p. 31)。つまり、どれだけ引き出す方法 (問題例) が具体的に論じられたとしても、そこで想定されている科学的探究の定義と結び付けて論じられな

れば、その評価の内実を十分に捉えたことにはならない。

そこで本稿では、NAEPのFrameworkについて、その背後にある科学的探究の議論に言及しつつ、その実際を明らかにすることを目的とする。それに際し、まずNAEPの科学の評価の位置づけを確認しておこう。

2. アメリカでのNAEPの位置づけと評価の枠組み

①アメリカにおけるNAEPの概要と位置づけ

アメリカでNAEPが初めて行われたのは、1969年まで遡る。NAEPは、個人の学習の到達度を判断するものではなく、アメリカの子どもたち全体の教育達成度を測定することを意図している。それにより、地域や国の教育政策を立案したり、既に施行された教育政策を評価したりするという形で政策的な判断を下す。

特に、2000年代以降には全ての州においてNAEP調査が実施されている。その背景には、2002年に制定された「どの子ども置き去りにしない法 (No Child Left Behind Act)」などで説明責任が厳格化されたことがある (吉良, 2012)。この時期には、NAEPでも州ごとの学力の経年的なデータの検出に促されるように修正が加えられた。こういった状況の下で「州の説明責任を果たすツール」という機能が付与されたNAEPに全ての州が参加していくようになる (木村, 2008)。その結果として、NAEPは、アメリカでも最大規模で、かつ影響力をもった学力調査の1つとなってきた。

NEAPには、主に3つの調査がある。まず、(A)長期にわたる動向評価 (long-term trend assessment)、次に(B)主調査 (main NAEP)、最後に希望する州のみで調査の精度を上げるために実施される(C)州別調査の3つである。これらは、それぞれ固有の目的で、異なる教科・周期・問題で実施される (NCES, 2013)。

全米規模で実施される調査は、(A)と(B)の調査である。(A)の動向評価では、学力の変化を長期にわたって追跡調査するため、英語や数学といった教科について同じ問題を使って繰り返し調査をする。近年では、4年周期で第4・8・12学年において実施されている。

他方で、本稿で中心的に分析していく調査が、(B)の主調査である。主調査は、かつてコールマン・レポートがそうであったように、教育条件の差や教育環境による教育到達度への影響を検出することが目的であった。近年では政策立案に必要な情報を社会から収集することが目的として掲げられ、英語や数学・科学、合衆国史や公民・地理といった科目で第4・8・12

学年において実施されている。ただし、年度によって実施される学年や教科にばらつきがある。

主調査の実施に向けて、各教科で作成されるものが、各教科の評価の枠組みである。この評価の枠組みは、先述のように NAEP において、何をどのような評価方法で問うのかということを示し、評価全体の設計の枠組みを提示する。そのため、評価の枠組みは、NAEP の評価の設計の中核を占めているものだといえる。そこで次に、この評価の枠組みについてみていこう。

② NAEP の評価の枠組み

NAEP の調査の枠組みは、一般の人々、そしてテストの開発者を主な読者としている。評価の枠組みは、テストの開発者が作問をする上での指針となるだけではない。一般の人々に向けて、テストの得点の背後にあるものを可視化することで、得点以上の意味、つまり、どのような力を子どもたちが身につけているのかということを広く伝えることが意図されている。

どのようにこの評価の枠組みが開発されるのだろうか。まず、NAEP の方針を決める母体である全米調査統括委員会 (National Assessment Governing Board) が、その教科領域の研究者を含む専門家集団に『討議報告書 (issues paper)』の作成を依頼する。この討議報告書が評価の枠組みの議論の土台となる。

討議報告書は、評価の計画・運営に関わる委員会が行う、その領域で評価を考える上で、重大な懸念事項についての討議を枠づけする手助けとなる重要な問いを議論の俎上へと挙げるものである。Framework に向けた議論では、「Framework に含まれる内容は、どのような広さであるべきで、広さと深さの選択への示唆はどのようなものであるか」(Champagne et al., 2004, p. 5) という評価対象に関する問いが提示されている。

加えて、討議報告書では、その課題と密接に関連した近年の評価等に関する研究成果についても報告されている。例えば、先の Framework では、国際的な学力調査の 1 つである PISA (Programme for International Student Assessment) や TIMSS (Trends in International Mathematics and Science Study) などのテストの枠組みについて、アメリカ国内で設定されている目標との類似点と相違点を明らかにした研究が紹介されている (Champagne et al., 2004, p. 5)。

この討議報告書にもとづいて評価の枠組みの草稿が作成される。これは、NAEP の請負業者と専門家団体によって行われる。ここで完成した草稿は、評価の専門家や教師や政策立案者、一般の人々など様々なス

テークホルダーに会議や学会の場などで公開され、議論され、省察され、改訂される (NAGB, 2008)。最後に提案された評価の枠組みは NAEP のウェブサイトに入れられ、公開ヒアリングが実施される。多様な人々の声や教科領域の研究成果等を反映できるように、完成までには約 2 年の歳月を要しており、重厚な議論の下で評価の枠組みが開発されているといえるだろう。

評価の枠組みの内容や構成は、教科によってばらつきはあるものの、多くの場合に次の要素が含まれている。(a)評価される内容とスキル、(b)学年ごとの到達点の目安、(c)設問領域の配分、(d)出題形式、(e)例示となる設問 (illustrative item) と(f)その採点指針がそれぞれ、理論的な根拠を伴って示される。

評価の枠組みは、テストごとに開発されるのではなく、経年での結果の比較や傾向が捉えられるように、一定の期間ごとに (少なくとも開発から 10 年以上は間を置いて) 改訂される (Jago, 2009, p. 1)。しかしながら、それは必ずしもその間は固定された問題に完全に縛られるということの意味するものではない。あくまで、評価の枠組みに沿って問題が新たに開発されるということも視野に入れられている。

科学教育については、先述のように 2009 年の実施に向けて開発された Framework が最新のものである。これは、2009 年・2015 年の調査で活用されている。今後も、2019 年・2023 年まで同 Framework が利用され、2027 年に必要に応じて新しい評価の枠組みが開発され、調査が実施される予定である (NAGB, 2019)。ここで科学教育において、どのような研究知見が反映され、何が論点とされてきたのか。科学的探究力の部分を中心に Framework の開発に向けた議論をみていこう。

3. Framework の開発に向けた議論とその内実

① 討議報告書と科学的探究をめぐる議論

Framework の開発に向け、2004 年に討議報告書が出版された。議長は、アメリカの認知心理学者であり、構成主義に関する研究者としても知られているシャンパーニュ (Champagne, A. B.) が務めている。その他にも、1995 年に発表された科学教育の共通教育目標である『全米科学教育スタンダード (National Science Education Standards)』の議長を担ったバイビー (Bybee, R. W.) や 2007 年に公表された全米研究評議会の報告書である『学校に科学を (Taking Science to School)』を取りまとめたデュース (Duschl, R.) などが識者とし

て討議報告書の表紙に名前を連ねている (Champagne et al., 2004)。

この報告書は、Framework の開発に向けて、全部で 18 の論題 (issues) と勧告を示している。具体的には、Framework の形態や他の国内外の学力調査との兼ね合い、素朴概念を考慮に入れた子どもたちの重大な観念 (big ideas) の理解の深まりへの配慮、全米科学教育スタンダード等の教科の専門団体が作成したスタンダードへの着目など幅広い内容が、新たな研究知見とともに提言されている (Champagne et al., 2004)。

この討議報告書において、科学的探究は資料 1 のようにまとめられている。ここで、科学的探究は自然科学という学問の中核に置かれるとともに、科学的リテラシーの重要な要素として位置づけられている。特に、ここで科学的探究が単なるプロセス・スキルの集合体として捉えられるのではなく、そこに内包されている認識論も含めて科学的探究が描かれている点に、この討議書の 1 つの特徴を看取できる。それは、「自然界を調査するために科学者が用いる方法として探究を理解する」という方法だけの理解に留まらず、「科学者によって実践されているように探究を理解する」(Champagne et al., 2004, p. 16) ことが報告書に掲げられている点にも端的に現れている。

この点について、Framework の科学的探究の特徴をより正確に描き出すために、討議報告書作成時に用いられていた枠組みと比較してみよう。1996年に発表された NAEP の科学の評価の枠組みでは、科学的探究という用語は用いられていない。それに代わり、「知ることと行うこと (knowing and doing)」という認知に関する次元が生物や物理などの内容に関する次元と並んで設定されている。これは、プロセス・スキルや認知に関するスキルを規定するものとされ、その中には、概念的理解、科学的調査、実践的推論という 3 つのカテゴリーが含まれている (NAGB, 1996, pp. 20-

資料 1. 討議報告書における科学的探究の位置づけ

探究は科学的リテラシーの重要な構成要素である……探究は、複雑で多角的な概念である。探究は、自然科学という学問を特徴づけるものである。探究は、自然界を調査するために科学者によって用いられている方法を内包している……自然科学の研究者によって実践されているように、探究はプロセスというよりも、むしろその哲学的な基礎 (特に、科学的知識を構成するものについての認識論や知見) によって特徴づけられるものである。

(Champagne et al., 2004, p. 16 を筆者訳出。なお、下線部は引用者による)

21)。

この科学的調査のカテゴリーは、それ以前に設定されていた「探究の実施」のカテゴリーに代わるものとして位置づけられている。このカテゴリーでは、科学的調査は「単なる科学の方法の別名ではない」と明確に記載されている (NAGB, 1996, p. 23)。そこでは、論文等で報告される科学の在り方と科学者が行う科学の間の乖離の問題が取り沙汰されている。ここで、科学的調査は、論文にあるような単純化され、直線化された営みではなく、現実の科学者の営みへと近づけていくことを志向して設定されているものであることがわかる。

ここで指摘されている現実の科学とはどのような営みを指すのだろうか。評価の枠組みでは、科学者の仕事の中心は公正な検証、つまり、条件統制された実験にあると記されている (NAGB, 1996, p. 24)。その上で、それを含む科学的調査とは「新しい情報を獲得し、適切な調査を計画し、様々な科学的な道具を用いて、その調査の結果を伝達する」(NAGB, 1996, p. 21) こととして描かれている。ここでは、複数のプロセスや道具を組み合わせて、公正に検証する営みとして科学的調査が描かれている。この定義を受けて同カテゴリーでは「認知に関する、もしくは実験室の道具の両方を含む科学の道具を使う生徒の能力を明らかにする」(NAGB, 1996, p. 21) ための評価方法の開発が意図されている。

先の討議報告書の探究に関する定義と比較してみると、1996年の評価の枠組みの科学的調査に関する記述には、厳密に科学的探究のプロセス、特に実験活動を遂行する以上のことが十分に内包されていないといえる。もちろん、1996年の評価の枠組みにおいて、2004年時に見られたような認識的な要素が全く含まれていないわけではない。だがそれは、科学の本質 (nature of science) という形で通常の内容や手続きに関する次元とは切り離された区分に設定されていた (NAGB, 1996, pp. 26-27)。それが、討議報告書の科学的探究の定義には含意されていたという点に違いがみられる。

それを踏まえて、Framework の開発に向けてどのような論題と勧告が提言されたのだろうか。資料 2 は、科学的探究に関わる主な論題と勧告をまとめたものである。ここからは、科学的探究に関して測定すべき要素を特定することが第 1 に掲げられ、それに付随して、認知に関わる側面や科学的探究の本質に関わる内

資料2. 討議報告書の科学的探究に関する主な論題と勧告

論題	勧告
([科学的]) 概念と同様に複雑な ([科学的]) 探究は、測定される構成概念としての明確な仕様書を開発することに対する挑戦を提示している。	Framework は、評価において測定される探究の要素について明確に伝達しなければならない。
どんな科学的探究の見方が Framework に含まれるべきだろうか。	委員会のメンバーの Framework における探究に関する決定は、アメリカの全米もしくは州のスタンダード、もしくは TIMSS や PISA の枠組みに含まれる探究から情報を与えられるべきである。
Framework の一部となるべき認知に関する次元はどれだろうか。	児童・生徒のテストのパフォーマンスの解釈は、認知に関するモデルの点で自然科学の学問を横断的に行われる。探究の能力は、評価から得られる情報に有用な次元を加え、Framework に含めるために熟慮されるべきである。
科学的探究の本質は人間が学習する方法を学習し続けることで発展し続けていくだろう。科学的探究の規範的な特徴は純粋な実験を超えて拡張され、今では理論やモデルの構築と改訂を含んでいる。	Framework の委員会は、科学的探究の本質の発展に関する研究に熟知するようになるべきであり、Framework のための示唆を検討するべきである。

(Champagne et al., 2004, pp. 16-17および p. 19をもとに筆者作成。なお、下線部は引用者による)

容を検討することが企図されていることがわかる。

更に、先の資料1の議論についても勧告に反映されている。下線部で述べられているように、単に純粋な実験を科学的探究として取り入れるのではなく、理論やモデルの構築も含めたより広い営みとして科学的探究を位置づけていくことが示唆されている。そして、それを通して、科学的探究ということを認識論的に把握していくことが勧告されているといえよう。

では、これらの勧告はどのように Framework に引き取られ、解消されてきたのだろうか。Framework に即して検討してみよう。

② Framework における科学的探究の位置づけ

Framework は、それまでの NAEP の評価の枠組みと同様に科学の内容と科学の実践 (science practice) という2つの次元から構成されている。科学の内容は、NAEP で扱われる学問領域を指し、これまでの NAEP の評価の枠組みと同様に物理科学・生命科学、そして

地球と宇宙の科学の3つの領域から成る。また、形態と機能 (form and function) といった複数の領域に共通の内容として領域横断的内容 (crosscutting content) も置かれているが、これを含めても特段、既存の内容と比べて新しい内容は設定されていない¹⁾。

他方で、この科学の実践という次元は、Framework において初めて登場したものである。これは、NAEP の運営委員会 (steering committee) が先の討議報告書の勧告を受け、全米科学教育スタンダードなどを精査した結果として「Framework は、科学の本質と実践を反映する」べきという指針を提起したことを受けて記述されたものである (NAGB, 2008, p. 5)。

この科学の実践は、次の4つの実践から構成されている。まず①科学の原理の特定、次に②科学の原理の使用、そして③科学的探究の使用、最後に④工学に関する設計の使用である (NAGB, 2008, p. 66)。ここで、科学的探究は科学の実践の1つの構成要素として位置づけられていることがわかる。

Framework では、①から④の個々の実践の中身は具体的に示されているものの、科学の実践それ自体がどのような営みなのかということは明確に示されていない。この実践という考え方は、ほぼ同時期に先のデュースが中心となって取りまとめた『学校に科学を (Taking Science to School)』(2007) に登場するものである。この報告書は、新しいスタンダードの開発に向けて編纂され、「実践として科学を教える」という理念を全米規模に広げる1つの契機となったものである。そこで『学校に科学を』を頼りに、科学の実践の理念や内実についてみていこう。

『学校に科学を』では、多くの教室で教えられている科学の方法を次のように批判する。「長年 [教室で] 教えられてきた古典的な科学の方法は、科学者の営みの一般的な類似品を提供するものである」(NRC, 2007, p. 27) と。ここでは、1996年の評価の枠組みと同様に、教室における科学の教授と科学者の行う科学の営みの間の断絶が指摘されている。

その上で、同報告書では、「現在の科学教育は……実験を過度に強調する傾向にある。実験は科学における調査の一つの基本的な形態であるものの、それは決して単一の限定的な手段ではない」(NRC, 2007, p. 342) と記されている。ここで、討議報告書の議論と同じく、多くの教室において科学的探究をすることが、条件統制をすることなどデータの収集とその分析の手続きといった実験的な活動に矮小化されがちであ

ることが批判される (NRC, 2007, p. 342 など)。

そこで『学校に科学を』では、科学とは単なる実験より、幅広い実践へと従事するものであるというビジョンを提示する。そこでは「科学は、理論とモデルの構築や、そのモデルや理論の内的な一貫性や整合性の確認、そして実験を通じた検証というプロセスであり、そのように提示されるべきである」と綴られている (NRC, 2007, p. 342)。つまり、科学者の営みとしてこれまであまり光が当てられてこなかった理論構築に向けたモデル化やアーギュメントといった側面にまで、実践という視点から照射することを意図している。

ここで科学の実践に参加するということには、これを反映して2つの意味が込められている。1つが、科学的探究に関わる個々のプロセス・スキルの教授を行うことからの脱却である。そして、もう1つが科学における理論と実験の往還関係を強調することである。

全米科学教育スタンダードにおいて「探究とは、観察・推論・実験といったスキルを児童・生徒が学習するような『1つのプロセスとしての科学』を超えたステップ」であり、「いくつかのプロセスと科学的知識を結合することが求められる」(NRC, 1995, p. 105)とされている。これは、Framework における「探究が科学の本質の理解を含むために『プロセスとしての科学』を超えていくこと」(NAGB, 2008, p. 73)という文言の中に引き継がれている。両者は、それまでアメリカで隆盛していた文脈に依らないプロセス・スキルの個別的な教授からの脱却を目指していることがこれらの文言からうかがえる。

『学校に科学を』においても、科学の実践に参加することは「個々のスキルを別々に教え、児童・生徒にそれらの練習をさせるというよりも、むしろ必要に応じて彼女・彼らによって開発された問いに関連するより大きな調査の文脈において教えられるはずである」(NRC, 2007, p. 255)と述べられている。ここでは、「プロセスとしての科学」にあるように脱文脈化された状況で個々のスキルを教えるというよりは、むしろ、より大きな文脈の中で科学を教授するという意味を含んでいる。その結果として、子どもたちが科学の実践を通して複数の知識やプロセスを結びつけて、現実世界のように複雑な文脈で使えるようになることが企図されている。

他方で『学校に科学を』では、科学の実践においてアーギュメントの過程や説明の構成、モデルの構築と

いった、これまで実験の影に隠れ、あまり焦点化されてこなかった側面に子どもたちを参加させる機会を与えることを求める。このような実践に参加する際には、理論をもとにモデルを作ったり、結果をもとに説明を改訂したりするなど理論と実験活動が密接に結びつけられている。加えて、そこでは子どもたちが、科学者共同体が保持している学問の慣習ともいえる特定の言説を用いて、自分や他者の考えについて提案したり、アーギュメントの過程に参加したりする。このような討議の過程を経て合意に達するという、まさに社会的な営みとして科学者の営みを捉え、それを科学の実践へと反映している (NRC, 2007, p. 258)。

これは、科学社会学者のピッカリング (Pickering, A.) が提唱する科学的実践論にも通じる。科学的実践論では、まず科学の理論形成において論理的な関係性のみを問題とする表象主義の立場を批判する。その上で、世界と精神の間にある実験室の装置や科学者の行為などに含まれる物質性や行為遂行性の概念を科学教育へと持ち込む。具体的には、学術的な場において理論が受け入れられず反証されるといった営みを繰り返す中で、より理論が精緻化されていくといった、人間や対象、組織や社会関係からの抵抗と適応のプロセスの中で理論形成や方法の吟味がなされることを科学の営みとして重視するという立場をとる (Pickering, 1995)。科学の実践も、このような考えを反映し、理論と実験が社会的な場で複雑に絡み合いながら科学が形成されていくものとして捉えているといえる。

以上から、この科学の実践という視点は、必ずしもこれまでの科学的探究の概念と相反するものではない。それは、Framework において、科学の実践に科学的探究が含まれていることから推察される。つまり、上記のように実験活動に矮小化されがちな従来の科学的探究の概念を社会的側面も含めて拡張するという点で、科学の実践と科学的探究は、ある種の包含関係にある。加えて、知識や個々の実践同士の結びつきを強調することで、これまでの現実の科学者の営みに近づけていくという議論を継承・発展させて、提案されたものであったといえるだろう。

具体的に個々の実践についてみてみよう。資料3は先の①から④の科学の実践の概要について一覧にまとめたものである。これらの科学の実践は、先の内容に関する次元の知識をどのように使うのかということをも明確にするものとして位置づけられている。

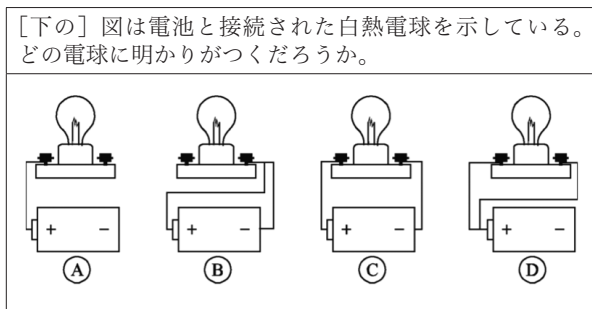
これらの4つの実践は、必ずしも段階として明瞭に

資料3. 科学の実践の概要

科学の実践	概要
科学の原理の特定	<ul style="list-style-type: none"> ・ 観察結果を記述・測定・分類する ・ 正しい科学の原理を叙述・認識する ・ 密接に関連した科学の原理の間の関係を示す ・ 科学の原理の様々な表現の間の関係を示す
科学の原理の使用	<ul style="list-style-type: none"> ・ 現象の観察結果を説明する ・ 現象の観察結果を予想する ・ 科学の原理を例示する観察結果の例を提示する ・ 代替的な説明や予想を提案・分析・または評価する
科学的探究の使用	<ul style="list-style-type: none"> ・ 科学的調査の側面を設計または批評する ・ 適切な道具や技法を用いて科学的調査を実施する ・ データのパターンを特定または、それを理論的なモデルと関連づける ・ 説明や予想についての結論が正当であると立証もしくは、批評するために実験にもとづく証拠を使用する
工学に関する設計の使用	<ul style="list-style-type: none"> ・ 規準と科学的な制約が課された問題に対する解決策を提案または批評する ・ 設計に関する決定において科学に関するトレードオフの関係にあるものを特定し、代替的な解決策の間で選択する ・ 技術的な設計に関する決定の影響を予測するために科学の原理やデータを応用する

(NAGB, 2008, p. 80をもとに筆者作成)

資料4. NAEP の例示となる設問



(NAGB, 2008, p. 70を筆者訳出)

区別されうるものではない。むしろ、これらは連続体として存在している。例えば、Framework では資料4の問題を例に挙げて、このことを説明している。この設問は、一見すると回路に関する知識を適用するだけなので、子どもの用いる科学の実践は、科学の原理を特定することにあると捉えられるかもしれない。しかし、回路そのものの学習が既習でなければ、そもそも推論に向けて原理を特定した上で、応用することが求

められる。その場合には、科学の原理の使用に属するものだとされる。ここで両者の関係は段階関係にあるのではなく、むしろ包含関係にあるといえる。

実際に作問する際には、資料4のように科学の内容に関する次元（物理科学など）と科学の実践に関する次元（科学の原理の特定など）を組み合わせるような問題を構想する。Framework では「科学の実践は内容に依存しないスキルではな」く、科学の内容に関する次元の知識や科学の本質、科学的探究に関わる知識を要するものだと述べられている (NAGB, 2008, p. 82)。そこで、それらを接合した形で問題を設定するために、両者の交点において記述されるものが、「期待されるパフォーマンス (performance expectation)」である。

この期待されるパフォーマンスは、学年段階に適切な科学の内容と科学の実践を組み合わせることで生成される。第8学年の物理科学を例にとれば、「[例えば、おもちゃの車などの] 物体は直線上を一定の速度で動いている。この物体が下り坂を転がるときに物体の速度はどのようになるのか予想することができる」(NAGB, 2008, p. 83) という期待されるパフォーマンスが設定される。これは、第8学年の物理科学の力に関する理解（例えば、物体の運動する方向・速度と物体に働く力の関係の理解）と科学の原理の使用という科学の実践が組み合わせられ、記述される。

ここで、学年段階に適切な形で設定する手助けとなるものが、ラーニング・プログレッションズ (learning progressions : 以下、LPs) である。LPs とは、適切な教授により子どもが長期間にわたり学習する領域の知識等を発達させる際に経る段階とその発達の道筋について仮説的なモデルを示す資料である (Corcoran, Mosher, & Rogat, 2009, p. 37)。これは、子どもの素朴概念や学問の論理に関する研究に立脚して、子どもたちが長期的なスパンで概念を発達させていく際に経る道筋とその時々到達点を示したものである。

先の物理科学を例にみてみよう。物理科学の運動に関して NAEP で構想された LPs では、まず第4学年で物体の動く変化と力の関係を学ぶ。次に第8学年では、力が向きと大きさをもった量（ベクトル量）であることや磁力・重力など離れた場所で作用する力が存在していることを知る。加えて、等速直線運動など力が外的に作用しないときに物体に生じる運動などに関する定性的な知見を得る。最後に第12学年では、これらの力と運動の関係を量的に捉える。運動方程式

($F=ma$) などはその一例といえる (NAGB, 2008)。このように質から量へ、力が捉えやすいものからより抽象的なものへと深まっていくことが想定されている。

NAEP の全ての評価項目は、この LPs に裏付けられた期待されるパフォーマンスから引き出される。主な評価方法としては、筆記テストとハンズ・オンの課題、そして双方向のコンピュータの課題 (interactive computer task) がある。また、回答形式としても多肢選択と自由記述 (概念地図も含む) を含んでいる。

4. NAEP における評価方法の実際

具体的にどのような評価が行われているのだろうか。資料 5 は、2009 年および 2011 年に行われた第 8 学年の NAEP の物理科学の問題の筆記テストの一例である。それぞれの問題で、先の期待されるパフォーマンスに沿った課題が設定されている。例えば、科学の原理の特定では、水分子の組成を記憶から再生する力を問うている。これは、「元素は 1 種類の原子から構成される純物質を構成している。化合物は、2 つ以上の異なる元素から構成されている」(NAGB, 2008, p. 33) という物理科学の内容と結びつけられている。

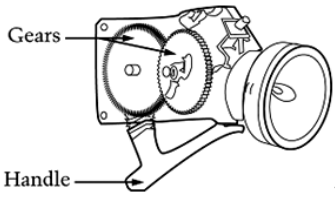
他方で、科学の原理の使用では、周期表に含まれる周期律の考え方を軸に類似の性質をもった元素を予想する力を測っている。この問いは、「すべての純物質は、約 100 個の元素のうち 1 個以上の元素で構成されている。周期表は、類似の特性をもつ元素を族へと編成するものである」(NAGB, 2008, p. 33) という物理科学の内容を使う問題として設定されている。

先節で確認してきたように、科学の実践という場合には、プロセスや知識の間のつながり、そして批評という要素が強調されている。このような要素を含んだ評価の問題に本稿で中心的に検討している科学的探究の使用に関する設問がある。

この科学的探究の使用の設問では、例えば、資料 5 のように水が水蒸気に変化する物理的变化と木片が炭化する化学変化を見た子どもが、両者を化学変化であると結論づけたことに対して、科学の原理に立脚して批評することが求められる。そこでは、2 つの物質の変化の違いを原理として特定した上で、証拠にもとづいて自らの論を論証する。

以下は、この設問に対する生徒の回答例である。ここでは、「木片が燃えたことは、化学変化である。なぜなら、その変化は新しいものを生み出しているからである。水が蒸気になったことは、物理的な変化で

資料 5. NAEP の物理科学における調査問題例

実践	問いの例
科学の原理の特定	<p>どんな原子が水の分子を構成しているか。</p> <p>A. 1 個の水素と 1 個の酸素 B. 1 個の水素と 2 個の酸素 C. 2 個の水素と 1 個の酸素 D. 2 個の水素と 2 個の酸素</p> <p>(2011 年に実施)</p>
科学の原理の使用	<p>[周期表が与えられている]</p> <p>上記の周期表での [各元素の] 位置にもとづいて、どの元素がアルゴン (Ar) と最も類似した科学的な特性を持っているか、予想しなさい。</p> <p>A. 塩素 (Cl) B. ヘリウム (He) C. 窒素 (N) D. 亜鉛 (Zn)</p> <p>(2011 年に実施)</p>
科学的探究の使用	<p>授業で次の 2 つの演示実験を観察した。まず、蒸気へと変わる水を、次に木片が燃えて、煙を出す様子を観察した。生徒は両方の演示実験は化学変化の例であるに違いないと結論づけた。なぜなら、それぞれの実験で気体が生成されたからである。この生徒の結論は正しいだろうか。両方の演示実験に言及しながら、あなたの回答を説明しなさい。</p> <p>(2009 年に実施)</p>
工学に関する設計の使用	<p>下記の懐中電灯には電池がない。ハンドルを握って離すことによって操作される。懐中電灯の本体の内側には歯車 (Gears) があり、それを以下に示している。</p>  <p>どの順番で書かれたものが、光を生み出すために懐中電灯の中で行われるエネルギー変換を最もよく特定しているか。</p> <p>A. 運動 → 電気 → 光 B. 運動 → 化学 → 光 C. 化学 → 運動 → 光 D. 化学 → 電気 → 光</p> <p>(2009 年に実施)</p>

(<https://nces.ed.gov/NationsReportCard/nqtf/> 2019.11.30 確認より筆者訳出)

ある。なぜなら、蒸気は水へと戻ることができるからである」(NAEP Question tool) と記されている。

ここで、この回答をした生徒は、まず化学変化と物理的变化という必要な知識を特定し、それをもとに自分なりに現象を説明している。その上で、気体の発生の有無は化学変化と物理変化を特定する要因にはならず、実際には生成物の有無やその性質の変化、可逆変化か不可逆変化かという点が検討されることを通して両者は特定されるという観点から、もとの結論を批

資料6. 知識の種類と認知に関する要求

知識の種類	認知に関する要求
概要	
1. 宣言的知識	Knowing that
基本的な科学の事実や概念や原理を知り、推論できるとともに、適切にこれらの基本原理を再生し、定義づけし、提示し、使用し、関連づけること	
2. 手続き的知識	Knowing how
科学するときに事実や概念や原理を応用すること	
3. 概略的 (Schematic) 知識	Knowing why
自然現象を予想・説明し、また、なぜ、どのように科学的主張が評価され、論じられ、正当化され、反証されるのかということを説明すること	
4. 方略的知識	Knowing when and where to apply the knowledge
自身の現在の知識を取り上げ、それをどこか新しい状況に適用すること	

(NAGB, 2008, pp. 91-92をもとに筆者作成)

判している。このように科学的探究に関する設問では証拠をもとに、他の結論を批判的に吟味し、科学的知識を用いて解決する力を問おうとしているといえる。

特に NAEP では、このような問題に回答する際に実践において働く知識の種類を特定することで、その活動で、どんな認知に関する要求 (cognitive demand) があるのかということと同定している。Framework によれば、この認知に関する要求には、資料6の4種類の知識に対応して、4種類の要求があるという。

それぞれの知識は、科学の実践と密接な関係性を有している (資料7)。例えば、宣言的知識に結びつく認知に関する要求とは、科学の原理を特定することと関係が深い。それは、資料5において科学の原理の特定が水分子の組成を想起し、提示するという課題と結びついていることからわかるだろう。

資料7では、科学的探究の使用が最も多くの知識と結びつきが強いとされている。これは、科学的探究を行うことが、プロセスに関する知識だけでなく、それ

資料7. 認知に関する要求と科学の実践との関連性

	原理の 特定	原理の 使用	探究の 使用	工学の 使用
宣言的知識	○			
手続き的知識			○	
概略的知識		○	○	○
方略的知識		○	○	○

※○は特に結びつきが強いことを示す

(NAGB, 2008, pp. 91-92をもとに筆者作成)

をどのような場面で用いるのか、更にはなぜそれを用いなければならないのかというメタ的な知識があって初めて成立するものであるということを示しているといえる。

ここで資料7を見ると、宣言的知識は原理の特定にしか作用しないと誤解されるかもしれない。しかし、「認知に関する要求は、(科学の実践と同様に) 関連していて独立していない」(NAGB, 2008, p. 92) と Framework に記されているように、それは正しい理解ではない。例えば、概略的知識を用いて科学的主張の正当性を論じる場面では、宣言的知識を想起する必要があることは容易に想像できる。

あくまで、資料7は、個々の実践について主に作用する認知過程を示しており、それらは総合されて現実には適用される²⁾。科学の実践において原理の使用と特定が包含関係にあったように、現実場面に知識を適用する、つまりいつ、どこで知識を適用するのか知ることは、そもそもある知識を知ることが内包されている。その上で、子どもたちのパフォーマンスを解釈する主な参照軸となる枠組み (認知モデル) として、これらの知識を導入し、主たるパフォーマンスの背後にある認知過程を特定するのである。

このように Framework では、期待されるパフォーマンスで示されるように、科学の内容と科学の実践という2軸で評価項目の作成を試みていた。しかしながら、科学の実践を位置づける際に、それを単なる行為として捉えるのではなく、その中にある認知過程も含めた形で構想することで、より緻密に子どもたちの幅広い能力を測り、解釈しようとするものであった。

だが、科学的探究力の評価というときに、単なる知的操作という表象のレベルに留まり、実際に実験などで器具を操作したり、長期的な観測を行い、得られたデータを解釈したりする行為のレベルを含めなければ正当な評価ではないという反論があるかもしれない。確かに、科学の実践という概念を提起することで、実験活動に留まらない理論構築等に目を向けつつ、他方で元来の行為のレベルにおいて特にヒトやモノからの抵抗が捨象されてきたことが糾弾されてきたことに鑑みれば、その点を含みこまないことは片手落ちといわざるを得ない。

この点について、Framework では時間と資源の面から次のような活動で身につけた児童・生徒の学力を評価することの困難さが述べられている。具体的には、研究上の問いを枠づけ・推論すること、データ収集の計画とその遂行、実際の調査において生じた予期せぬ

挑戦を解決すること、新しい証拠に照らして自分たちの実験に関するアプローチを新しくする方法について決定する議論に参加することといった活動が挙げられている (NAGB, 2009, pp. 8-9)。

しかしながら、他方で Framework では、「測定することが困難で時間を消費するが、科学者や科学教育者や実業界によって価値づけられている、いくらかの科学教育の重要な成果 (例えば、知性の習慣や持続的な探究、協働的な研究) は、Framework や NAEP の科学の評価において部分的にはあるが提示されている」(p. 8) と記されている。NAEP の評価では、これまで見てきた筆記による評価以外にも、先行研究が指摘しているように様々な評価方法が用いられている³⁾。

例えば、2009年に行われた第8学年の物理科学のハンズ・オンの課題を見てみよう。まず、子どもには実験用の棒磁石と4本の金属棒、そして鋼製のワッシャーが入った袋と定規とグラフ紙が与えられる。子どもたちは、指示に従って課題を行うことが求められる。

最初の課題は、4本の金属棒がそれぞれ強い磁石、弱い磁石、銅棒、鋼棒のどれであるのかを特定する手続きを記述し、実施するとともに、そこから得られた観察結果を記録し、特定する。次の課題では、先の課題で得られた結果について、実験用の棒磁石を用いて検証する手続きを記述・実行し、観察結果を記録することで、再び特定を試みる。更に、確認の際にどのように実験用の磁石を用いたのかということを説明する。最後の課題では、これまで実験で明らかにされた強い磁石と弱い磁石について、与えられたすべてのものを用いて、その強さと弱さを示す実験を2つ以上計画し、実施した上で、その結果を記録する。例えば、ワッシャーを引き寄せる個数や鋼棒を引き寄せられる距離などが記録される。

これらの課題では、実際に物質を操作し、科学の内容や実践を駆使しながら問題を解決する。その中では、金属棒を特定したり、様々な磁石の長所を比較したりするための手続きを設計することや、観察結果を記録すること、そして調査にもとづいて説明を提供することなど幅広い内容や認知に関する要求が評価される。加えて、連続した実験の中で自らの実験結果を問い直したり、証拠にもとづいて結論を書いたり、修正したりする際に、「事実と意見が混在していたり、提示された証拠と結論が論理的に整合していなかったりするとき、議論は破綻する」などといった、批評する力や科学の本質に関わるような内容も同時に問われる。

この他、シミュレーションの課題はどうだろうか。第8学年では異なる液体を注ぐ際の液体の動きと温度の関係について調べる文脈が提示される。まず、最初の課題では温度による液体の性質の変化を扱う。例えば、メスシリンダーに入った同量のはちみつ、コーンシロップ、水、オリーブオイルについて、コンピュータ上で異なる温度に設定し、そこに銅球を落下させた時に銅球が底に到着するまでの時間を測定するというシミュレーションを行う。次に、同様の道具を用いて20°Cの時に最もゆっくりと流れる液体や、30°Cの時に水と同じ速度で流れる液体について選択肢から回答する。加えて、どのようにそれを知ることができたのかデータをもとに説明する課題も設定されている。

次に、様々な液体について低温の状態から高温になった時に早く流れる液体を調べる方法を考え、記述する。その上で、それを用いたシミュレーションをした結果、どの液体が早く流れたのか、それはなぜいえるのかということデータをもちぎ説明する。最後の課題として、食品加工会社が温めたはちみつを瓶に詰めている事例が挙げられ、できるだけ早く瓶に入れる方法について、何度か適切か、温度と流れる速度の割合のグラフを特定するとともに、そのグラフを使って瓶に入れるはちみつの温度について提案する。

これらの課題では、筆記テストによる評価とは異なり、物質を操作したり、与えられた生のデータを解釈したり、それをもとに試行錯誤したりすることが求められる。ここで急いで断っておくと、科学的探究の使用を評価するとしても、必ずしもハンズ・オンの課題やシミュレーションを伴うような双方向のコンピュータを用いた課題を用いるわけではない。先述のように科学的探究を行う力を問うとしても、データ表を提示し、どの結論がデータと整合性があるものなのかを尋ねたり、より信頼できるデータを生み出すための実験計画の在り方を提案させたりすることもある (NAGB, 2008, p. 74)。あくまで、標準化された筆記テスト以外の幅広い様式を認め、期待されるパフォーマンスに適した多様な評価方法を取ることで、到達度をよりよく把握することに主眼が置かれているといえよう。

もちろん、先に Framework で自ら言及していたように、NAEPにおいて自らで問いを設定し、それに向けて計画した実験を行う力や、知性の習慣、持続的に探究を行ったり、協働的に研究を遂行したりする力を問うことが難しいということは否定できない。ただし、この点についてアメリカの科学教育の評価に関す

る研究の文脈では、評価を次の3つのレベル、すなわち①教室レベル、②学校や学区レベル、③州レベルで区別し、③に属するNAEPのような大規模な評価と①に属する授業レベルの評価を繋ぐような議論が登場していることも確かである。

そこでは、①から③の評価を一貫させた形で、評価システムを構築していくことが提案されている。アメリカで科学教育の評価研究を行っているペルグリーノ(Pellegrino, J. W.)が指摘しているように「どのような単一のテストも生徒の能力について最も信頼のおける測定方法と考えられるものではなく、大規模な評価でも生徒を微細なレベルで理解する1つの道具となり(Pellegrino, 2012, pp. 101-102)、かつ指導や学習の改善に寄与するものになり得る(Pellegrino, 2014, p. 72)。あくまで、1つの評価ですべてを把握することはできず、それぞれの評価の特性や限界を自覚した上で、目的に合わせて組み合わせて、多面的に子どもを把握していくことが目指されているといえる。

それに鑑みれば、長期にわたって自らの設定した問いを持続的に探究していく力については教室レベルの評価を実施して、フィードバックを行うことで学習者を励まし、指導や学習の改善につなげていく。他方で、NAEPなどの大規模な調査は、個々の子どものためというよりも、カリキュラム全体を俯瞰的に捉えるために用いる。そこでは州や地域、もしくは学校において子どもたちの科学的探究力を育む取り組みを十分に支援できているのかということ問い直すとともに、教師が主観的に子どもを評価したり、自分の指導を絶対視したりすることから脱却し、子どもたちに1つの視点から自らの学習を改善していく手だてを提供する機能を果たすことが企図されうるだろう。

おわりに

本稿では、NAEPのFrameworkについて、その背後にある科学的探究の議論に言及しつつ、何をどのように評価しているのかという点について、その実際を明らかにすることを目的としてきた。

Frameworkでは、2009年以前の科学的探究の評価の枠組みが目指す方針と軌を一にし、現実世界で科学者が行う科学的探究へと評価内容を近づけることが企図されてきた。特に、Frameworkでは、それまで科学的調査と呼ばれていたカテゴリーを含んだ「知ることと行うこと」という次元に代わり、新たに科学の実践という次元が設定されていた。

この科学の実践という考え方は、2000年代後半からアメリカにおいて科学的探究に比してより広い考え方として提起されてきたものであった。それは、実験を中心としたプロセス・スキルの指導を問い直し、理論構築に向け、批評を軸とした理論と実践との往還関係や共同体による議論といった社会的側面の果たす役割を強調したものであった。加えて、様々な知識と実践の結びつきについても射程に入れられていた。それにより実際の科学に近づけ、そこで発揮される力を身につけているのかということ問い評価を目指していた。

Frameworkにおいても、この発想が反映されていた。そこでは、科学的探究は科学の実践の核として存在しているものの、あくまで1つの要素として位置づけられ、幅広い実践の営みが内包されていた。加えて、どのような領域の知識をどのような実践と用いるのかということを示した期待されるパフォーマンスに沿って問題が作成されるとともに、解釈において認知過程が分析されることで、行為と認知の両面から子どもたちの到達度を問う評価となることが意図されていた。

実際の評価においても筆記テストに留まらない様々な評価方法が活用されていた。それにより、表面上の実験の手続き等に関する知識を問うだけでなく、実際に実験計画し、物質的な抵抗を踏まえつつ、理論構築をしていくことをまでを含んだ幅広い科学的探究力を評価しようとしていた。

総括すると、NAEPにおける科学的探究の評価は、科学者の営みを反映するという理念の下で描かれた科学の実践の枠組みに立脚して、評価の構成概念として社会的側面までを含めた科学者の行為とそこに含まれる認知過程を詳らかにし、実際に知識やスキルを使って何ができるのかということを実験室等で使えるレベルで問うものであったといえよう。これは、知識と切り離された形で探究の行為や能力を質的に評価することとは異なる切り口の評価方法を提案するものであった。

もちろん、行論で指摘してきたように大規模な評価であるがゆえに限界を抱えていることも事実である。それを踏まえた上で、科学的探究力について教室レベルの評価などと大規模な評価をどのように接続していくことが望ましいのかといった評価の接続に関する問題を論じたり、アメリカにおける科学的探究の評価をめぐる論点に即してNAEPの科学的探究の評価のもつ意義や課題について言及したりするということに、本研究は十分に迫れていない。この点については、稿

を改めて論じたい。

注

* 愛知県立大学教育福祉学部講師

- 1) このような科学者に共通な見方に該当する部分を教授するという発想は、1960年代からアメリカの科学教育に見られるものである(詳細は、大貫(2018a)や大貫(2018b)を参照)。なお、NAEPの1996年の評価の枠組みにおいても、共通テーマという名称でほぼ同様の内容が、内容や方法の次元とは異なる部分に位置づけられていた(NAGB, 1996)。
- 2) Frameworkには、「1つ以上の認知に関する要求は、より複雑な科学の実践と結びつきうる。これらの結びつきは、児童・生徒が異なった学年レベルで評価課題に持ち込む知識によってシフトしうる」(NAGB, 2009, p. 92)とある。本文の例を挙げれば、回路についての知識に習熟していれば、資料4の問題は宣言的知識を問う問題になるものの、もしそれ自体に馴染みがなければ、新しい状況に知識を適用する方略的知識の問題としても解釈されうることになるであろう。
- 3) 以下は、Complete Task Libraryより引用した。

引用・参考文献一覧

荒井克弘・倉元直樹編(2008)『全国学力調査一日米比較研究—』金子書房。

Corcoran, T., Mosher, F. A., & Rogat, A. (2009) *Learning Progressions in Science: An Evidence-based Approach to Reform*, (https://repository.upenn.edu/cgi/viewcontent.cgi?article=1026&context=cpre_researchreports 2019.9.4確認)。

Complete Task Library (https://www.nationsreportcard.gov/science_2009/ict_tasks.asp 2019.12.3確認)。

Champagne, A., Bergin, K., Bybee, R. W., Duschl, R., & Gallagher, J. (2004) NAEP 2009 Science Framework Development: Issues and Recommendations (<https://www.nagb.gov/content/nagb/assets/documents/publications/reports-papers/frameworks/2009-science-framework-issues.pdf> 2019.12.3確認)。

Hanauer, D. I., Hatfull, G. F., & Jacobs-Sera, D. (2009) *Active Assessment: Assessing Scientific Inquiry*, NY: Springer.

石崎友規(2012)「小学校理科における問題解決能力の評価法の検討—NAEP ScienceにおけるGrade 4向けタスクを例に一」日本科学教育学会『日本科学教育学会研究会研究報告』27(3), pp. 11–14。

Jago, C. (2009) “A History of NAEP Assessment Frameworks”, *Paper Commissioned for the 20th Anniversary of the National Assessment Governing Board 1988–2008*, pp. 1–14.

木村拓也(2008)「2003年以降の全米学力調査の変質」荒

井克弘他編『全国学力調査一日米比較研究—』金子書房, pp. 178–203。

吉良直(2012)「アウトカム重視への政策転換—1989教育サミットから2002年NCLB法制定まで—」北野秋男他編『アメリカ教育改革の最前線—頂点への競争—』学術出版会, pp. 35–51。

NAGB (1996) *Science framework for the 1996 National Assessment of Educational Progress*, Washington, D.C.: the U.S. Government Printing Office.

—— (2008) *Science framework for the 2009 National Assessment of Educational Progress*, Washington, D.C.: the U.S. Government Printing Office.

—— (2019) “National Assessment of Educational Progress Schedule of Assessments”, (https://www.nagb.gov/content/nagb/assets/documents/naep/NAEP_assessment_schedule.pdf 2019.11.27確認)。

NCES (2013) “What Are the Differences Between Long-Term Trend NAEP and Main NAEP?”, (https://nces.ed.gov/nationsreportcard/about/ltt_main_diff.aspx 2019.11.27確認)。

NAEP Question tool (<https://nces.ed.gov/nationsreportcard/nqt/> 2019.12.3確認)

NRC (1995) *National Science Education Standards*, Washington, D.C.: National Academy Press. なお、全米研究評議会(長洲南海男監修)『全米科学教育スタンダード』(梓出版、2001年)も参照。

—— (2007) *Taking Science to School*, Washington, D.C.: The National Academy Press.

二宮衆一(2019)「探究学習における教育評価のあり方」日本教育方法学会『中等教育の課題に教育方法学はどう取り組むか』図書文化, pp. 50–66。

西岡加名恵(2016)『教科と総合学習のカリキュラム設計: パフォーマンス評価をどう活かすか』図書文化。

大久保貢・森幹男・中切正人(2018)『「探究力」に対するルーブリック評価の開発』『大学入試研究ジャーナル』28, pp. 53–59。

大貫守(2018a)「ロバート・カープラスの科学教育論に関する検討—SCISの取り組みに着目して—」『日本教科教育学会誌』2018年, pp. 95–105。

—— (2018b)「R. W. バイビーの5E指導モデルに関する検討—R. カープラスの学習サイクルとの比較を通して—」『京都大学大学院教育学研究科紀要』2018年, vol. 64, pp. 373–385。

Pellegrino, J. W. (2012) “The Design of an Assessment System Focused on Student Achievement: A Learning Sciences Perspective on Issues of Competence, Growth, and Measurement”, Bernholt, S., Neumann, K., & Nentwig, P. (Eds.) *Making it tangible: Learning outcomes in science education*,

- NY: Waxmann, pp. 79–107.
- Pellegrino, J. W. (2014) “Assessment as a positive influence on 21st century teaching and learning: A systems approach to progress”, *Psicologia Educativa*, 20, pp. 65–77.
- Pickering, A. (1995) *The mangle of practice*, IL: University of Chicago Press.
- SSH 連絡会『SSH 先進 8 校による「探究型学力 高大接続研究会」での取組 (2018年 3 月)』(https://docs.wixstatic.com/ugd/60eb20_a200ad7a6d8d4ff7a577b00625666b46.pdf 2019. 12.3 確認)
- 丹沢哲郎 (2012) 「アメリカにおける科学の学力の捉え方」日本理科教育学会編『今こそ理科の学力を問う』東洋館出版, pp. 34–39。
- 徳永聖一・坂本憲明 (2013) 「アメリカにおける科学的リテラシーの定義に関する基礎的研究—NSES および NAEP を参照として—」日本科学教育学会『日本科学教育学会研究会研究報告』28(2)、pp. 117–122。
- 安野史子 (2008) 「全米学力調査の問題設計」荒井克弘他編『全国学力調査—日米比較研究』金子書房, pp. 35–57。