福井大学リポジトリ
UNIVERSITY OF FUKUI REPOSITORY

G・ウィギンズの教育評価論における「真正'性」概念 —「真正の評価」論に対する批判を踏まえて—

メタデータ	言語: Japanese
	出版者:
	公開日: 2009-04-17
	キーワード (Ja):
	キーワード (En):
	作成者: 遠藤, 貴広
	メールアドレス:
	所属:
URL	http://hdl.handle.net/10098/1975

The Japanese Journal of Educational Objectives and Evaluation Studies, 13, 2003, 34−43

G.ウィギンズの教育評価論における「真正性」概念

——「真正の評価」論に対する批判を踏まえて——

京都大学大学院
遠藤貴広

はじめに

本稿は、「真正の評価 (authentic assessment)」論について、その提唱者の一人とされるグラント・ウィギンズ (Grant P. Wiggins: 1950-) の所論に即して考察するものである。

「真正の評価」論は、大人が現実世界で直面するような課題に取り組ませる中で評価することを志向するアプローチである。それは、1980年代後半以降の米国において、従来の「標準テスト (standardized tests)」に依存する体制を批判する中で模索された。日本では、「総合的な学習の時間」の導入を契機に、各教科で支配的であったペーパーテスト主導の評価観を疑問視する動きがより強まる中で、「真正の評価」論に衆目が集まった。そして、最近広く実践されるようになった「ポートフォリオ評価法 (portfolio assessment)」の理論的支柱を成しているのも「真正の評価」論である。

カミングとマクスウェル (Cumming & Maxwell, 1999)によれば、学習と評価の文脈で「真正の (authentic)」という言葉を初めて公式に用いたのはアーチボルドとニューマン (Archbald & Newmann, 1988) であり、「真正の評価」なる言葉を使い始めたのはウィギンズ (Wiggins, 1989a) である。しかし、教育界に広く「真正の評価」論が浸透する中で、アーチボルドとニューマンが唱えた基本原理は実際には持続されなくなっているという。澤田稔 (1997) が述べているように、「『真正の』という言葉は、『民主的』という言葉と全く同様に、誰にも反論の余地のない、理想的な意味内容を含むように見えるので、この大いなるスローガンの傘の下には、さまざまな社会的立場にある人々が雲集する」(p. 58) のである。

このような現状を踏まえれば、「真正の評価」とは何か、その基本原理を再確認するとともに、それがなぜ必要だったのか、その提唱理由を明確にする作業を、論者を絞った上で行う必要が出てくる。そこで、本稿では、「真正の評価」論の提唱者として米国で積極的な発言を続けているウィギンズの教育評価論に対象を絞り[1]、まず、彼が1980年代後半に「真正の評価」論として提唱した内容を概観する。そして、彼の「真正の評価」論を批判する論考を取り上げ、その批判の論点を踏まえて、最終的に、彼が教育評価論に導入した「真正性」概念の特質に迫りたい。本稿により、ウィギンズが「真正の評価」なるものを必要とした理由が明らかとなろう。

1.「真正の評価」論の提唱

ウィギンズの「真正の評価」論が提唱されたとされる1989年、「真正のテスト」[2]の一例として彼は次のような課題を紹介している[3] (Wiggins, 1989a, pp.706-707)。

【口述歴史プロジェクト（9年生）】

あなたは、インタビューや文書資料に基づいて口述歴史を完成させてから、授業中口頭で自分の知見を発表しなければなりません。題材の選択はあなたに任せます。例えば、自分の家族、小ビジネス経営、麻薬中毒、労働組合、十代の親、最近の移民といったトピックが考えられます。自分の予備調査に基づいて仮説を3つ立て、それぞれの仮説を検証するための質問を4つ考えて下さい。評価規準は以下の通りです。

- 3つの仮説について調査している。
- 長期間に渡る変化について少なくとも1つは述べている。
- 自分で背景調査をしたことを示している。
- 資料として適当な4人にインタビューしている。
- 各仮説に関連する質問を4つ以上準備している。
- 誘導や偏りのない質問をしている。
- 適切な時に追質問をしている。
- 得られた回答の中で事実と意見の重要な違いを示している。
- 最も良い仮説を支持するための証拠を用いている。
- 自分の書いたものと授業で発表したものをまとめている。

そして、このような「真正のテスト」に共通する基本特性をウィギンズは次のようにまとめている。「第一に、その分野でのパフォーマンス[4]を真に描写するよう設計されている。評定の信頼性やテストの計算法が問題とされるのは、このときだけである。第二に、評価に使われる規準の指導と学習に非常に大きな注意が払われている。第三に、従来のテストに比べて自己評価が果たす役割がたいへん大きい。第四に、自分の見かけの習得が本物であることを確かめるために、生徒は頻繁に自分の作品を示し、公に口頭で自分自身を弁護することを期待される」(Wiggins, 1989b, p.45)。彼はなぜこのような「テスト」を提起するに至ったのだろうか。

指導の役に立たない集団準拠 (norm-referenced) テストの数値が一人歩きし、統計的正確さや節約を名目に人の判断が無視されていた状況を憂えたウィギンズは、「本当のテストとは何か」と問う。この問いに対して、彼は、「知的能力についての本当のテストは、典型的な課題でのパフォーマンスを必要とする」(Wiggins, 1989a, p.703) とした上で、次のように述べている。「第一に、真正の評価は、作家、ビジネスマン、科学者、コミュニティー・リーダー、デザイナー、あるいは歴史家に典型的に差し迫る難題やパフォーマンスのスタンダードを模写している。そこには、エッセイやレポートを書く、個人やグループで調査を行う、提案や原寸模型を設計する、ポートフォリオを整理するなどといったことが含まれる。第二に、正当な評価は、個々の生徒や学校の文脈に敏感である。評価は、人の判断や対話を伴うとき最も正確で衡平な (equitable) ものになり、それで、テストされる者が問いを明確にすることを求めたり、自分の答えについて説明したりできる」(pp.703-704)。この第1点目の主張、つまり、大人が実際に直面する課題を模写するということが特に注目され、それが「真正の評価」論の出発点となった[5]。

このような主張の背景には、「テストが指導の中心になっている」というウィギンズの洞察があった。それは、「テストのために教える」という宿命を容認するものであり、それゆえ「テストは本物の知的挑戦を生徒に提供しなければならない」(p.704) という主張である。つまり、標準テストが悪影響を及ぼしているからと言って、テストを学習活動から遠ざけるのではなく、テストを、実施するに値するものに変えることで、そこに向けられる学習活動をも改善するという発想である。

このような（真正の）テストは、大きな労力を要し、時間のかかるものであり、さらに他人との比較が難しいとされる (p.704)。これが、それまでに似たようなものが提唱されながら定着しなかった理由であり、標準テストが根強く残る原因であった。しかし、いくら効率が良くても、それが学習や指導に役立つものでなければ意味がない。逆に、効率が悪く見えるものでも、それが改善に寄与するものであれば、実施する意味がある。ウィギンズは、その実施するに値するものを「真正の評価」に求めたのである。では、このときウィギンズが「真正の評価」に求めたものとは一体何だったのだろうか。

「真正の評価」の設計にあたって、まず「生徒に得意になってもらいたい実際のパフォーマンスが何であるかを最初に決めなければならない」ことをウィギンズは強調する (p.705)。したがって、評定方法などは、この「望ましいパフォーマンス」が決められた後に考慮される。それは、従来の標準テストに基づく評価が、効率化や序列化を志向するあまり、その子ができること、そして、その子がすべきことについて何も示してくれない状況を憂慮してのことである。テストは何より、目指すべきスタンダードを設定し示してくれるものでなければならない。

ここではまた、「知っていることの証拠」についての問い直しが図られている。ウィギンズにおいて「習得 (mastery)」とは、口先だけで答えることではなく、「思慮深い理解 (thoughtful understanding)」を伴うものである。そして、この「思慮深い理解」には、「問題や複雑な状況に、効果的な、変形力のある、あるいは斬新なことができる」という意味が込められている (p.705)。これを見ようと思えば、当然、現実に近い複雑な状況設定が必要となる。さらに、それが単なる偶然ではないということを保証するために、長期間に渡って、そして様々な文脈で観察する必要も出てくる。ウィギンズによれば、一般に生徒の短期再生 (short-term recall) について調べられることはあまりにも多いのに、生徒にとって最も重要な「精神の習慣 (habits of mind)」についてはほとんど調べられていないという。そして、この「習慣」を重視して、彼は次のように主張するのである。すなわち、「もし我々が本気で観念の思慮深い統制を生徒に示させようとするなら、単一のパフォーマンスでは不十分である。我々は、矢継ぎ早の質問への返答の中で吐き出される機械的な教理問答ではなく、生徒のレパートリーを観察する必要がある」(p.706)。

このような認識に立てば、いわゆる「目標に準拠した (criterion-referenced)」テストにも不適切なものが多くあることが分かる。というのは、その問題が人為的であり、その手がかりも人工的なものばかりだからである (p.706)。重

要なのは、そのテストが、大人が実際に直面する難題を模写しているかどうかである。そのため、複雑な状況の中で、曖昧にしか定義できない事象を扱うことになるのである。

以上のようなウィギンズの主張に対しては、批判もいくつか見られた。その批判とは一体どのようなものだったのだろうか。

2.「真正の評価」論批判の論点

「真正の評価」論に対しては、すでに様々な批判が寄せられているが、本節では、ウィギンズが直接対した論争のみを取り上げる。

1つ目は、サイザック（G. J. Cizek）による批判で、1991年に『ファイ・デルタ・カッパン（*Phi Delta Kappan*）』誌において展開された。これは、「真正の評価」と同一視されることの多い「パフォーマンス評価（performance assessment）」を批判するものである。

1990年代に入って推進が目立つようになったパフォーマンス評価に対して、サイザックは次のような批判を寄せている。すなわち、パフォーマンス評価と同様のものは昔からあり、新しいのは「パフォーマンス・アセスメント」という粋な響きのある言葉だけである。また、パフォーマンス評価推進者たちは、標準テストを否定するばかりで、その代替として自分たちが求めているものを何も示していない。さらに、彼らは、「表面的妥当性（face validity）」ばかりに目をやり、他の妥当性に関する技術的な問題を無視している（Cizek, 1991a）。これがサイザックによるパフォーマンス評価批判である。そして、これに対してウィギンズが以下のように反論した。

まず、「パフォーマンス評価の何がオリジナルなのか」という問いに対して、ウィギンズは、それは本質的な問いではないとし、それよりむしろ、「向上のための直接的なフィードバックを提供できない間接的な測定が教育を駆り立てているのはなぜか」というのが本質的な問いであるとする。また、妥当性については、その測定自体を目的にしないで、良いフィードバックを提供できるかどうかということを問題にすべきとする（Wiggins, 1991）。

このウィギンズの反論に、サイザックが反駁を加えることになる。そこでの論点は次のようなものである。まず、「テストのために教える」ということを容認できなかったサイザックは、評価の形式を変えさえすればそれでいい、とする動きには賛同できなかった。また、妥当性や信頼性は最低限でいいとする主張にも否定的で、もっと実質的なものでもって教育評価の質を高める努力をせよ、というのがサイザックの主張である（Cizek, 1991b）。

しかしながら、標準テストとパフォーマンス評価、双方に問題点があることは両論者が認めるところで、どちらか一方でいいとはしていない、という点で主張は共通している。これを踏まえれば、「多肢選択型かパフォーマンス型か」といった二項対立的な見方に限界があることは明らかで、別の座標軸が必要となる。それがウィギンズにおいては「真正性（authenticity）」なのである。

また、パフォーマンス評価における信頼性については、例えば「モデレーション（moderation）」[6]によって高められることが知られており（Wiggins, 1989a, pp.709-710）、サイザック

も「信頼できるパフォーマンス評価が現存する」ことを認めている（Cizek, 1991a, p.698)。したがって、ここでより大きな問題となるのは、後述する妥当性に対する認識の違いである。

2つ目に取り上げるのは、ターウィリガー（J. Terwilliger）による批判で、1997年から1998年にかけて『エデュケーショナル・リサーチャー（*Educational Researcher*)』誌において展開された。これは、ウィギンズの具体的な著作を対象に、「真正の評価」論を真っ向から否定するもので（Terwilliger, 1997)、この批判に対して、ウィギンズの他、ニューマンとブラントが反論し（Newmann, Brandt, & Wiggins, 1998)、さらに、ターウィリガーがウィギンズとニューマンに反駁する（Terwilliger, 1998）という形で展開された。

ここでウィギンズとの論争のみに限定した場合、ターウィリガーによる批判の論点は次の2つにまとめられる。それは、第一に、「真正の」という言葉の使用に関するものである。「本物の（genuine)」や「現実の（real)」という語と同様、「真正の」という語には誰にも否定しがたい理想的な響きがあるため、当時各方面で多用されていた。しかし、このような言葉が中立性を欠いていることは明らかで、そのような語を教育学という学術的な議論に持ち込むことを拒んだターウィリガーは、もし使うなら「代替（alternative)」という語にすべきだと主張したのである。第二に、ターウィリガーは、ウィギンズが知識の役割を軽視していることにも反感を抱いている。つまり、ウィギンズにおいては、学習の広さよりも深さが重視され、「きちんと構造化されない（ill-structured)」問題の中で働く「高次の思考（higher-order thinking)」が求められているが、そこでは知識が蔑ろにされているのではないか、という批判である。

以上2つの論争を踏まえて、次節では、妥当性や知識の捉え方にも触れながら、ウィギンズが教育評価に求めた「真正性」とは一体どのような概念であったのか、その特質を明らかにしたい。

3.「真正性」概念の特質

（1）現実的な文脈

ウィギンズは、テストの「真正性」を見る基準を挙げる中で、その特質を「真正のシミュレーション（authentic simulations)」という撞着語で要約している。そして、その追求に向けて、シミュレーションの「忠実性（fidelity)」（現実に近いか）と「包括性（comprehensiveness)」（多くの異なる側面を模写しているか）を高める必要があると言う（Wiggins, 1993, p.230)。そこで、テストにおける文脈の役割が問い直されることになる。

ウィギンズの「真正の評価」論においては、「パフォーマンスの多様で豊かな文脈を模写あるいはシミュレートする」ことが最も重視される。それは何より彼の次のような現実認識によるところが大きい。すなわち、「現実生活において、我々は、特定の問題を解決するために、特定の文脈の中で自分の知力や必要とされる知識や技能を用いる」（Wiggins, 1993, p.231）ということである。これを裏付ける知見として、例えば、状況論が援用される。そこでは、「認知は全て文化と文脈の『状況に埋め込まれ（situated)』ていて、脱文脈化した学習や評価は薄弱な機能不

全の――『有用な学習を十分に生み出すことのない』――ものとされている（Brown, Collins, & Duguid, 1989)」(Wiggins, 1993, p.232)。つまり、学校の中で行われる学習やテストは、学校の文化や文脈を反映したものであるため、学校内でうまくやれるからといって、それが学校外でうまくやれることを保証するわけではない。現実世界でうまくやれるかどうかを見ようと思えば、当然、現実世界の文脈が必要なのである。

ウィギンズは、このように文脈面を強調する中で、さらに、「制約（constraints)」の問題にも目を向ける。ウィギンズによれば、何かを実践するときに出てくる典型的な制約には次のようなものがあるという。すなわち、「自分自身の要求であろうとなかろうと、他人によって据えられた要求がある。課題を終えるのに使える時間に制限がある。状況や時間制限にもよるが、自分で思い通りにできる人的・物的リソースには制限がある。我々は進行しているので、得られる指導やフィードバックには制限がある」(Wiggins, 1993, p.234)。ウィギンズは、このような制約の存在を認める中で、従来の標準テストにおける制約があまりにも現実離れしていることを特に問題視した。そして、テストにおける真正性として、「問題を解決するのにたいてい有効である道具にアクセスできるようにしておいて、典型的に、そして『自然に』起こるような制約の下での質問や課題を生徒に経験させる」ことを要求したのである（p.236)。

したがって、ウィギンズの教育評価論における「真正性」とは、実際の文脈や制約を模写した現実的なシミュレーションの中で実現されるものだったと言うことができる[7]。

ここで、なぜ「真正の」という言葉を使わなければならなかったか、その理由に触れておきたい。まず、ターウィリガーが言うように「代替」という言葉を用いた場合、伝統的な標準テストに「替わる」ということは示せても、それがどのようなものであるかは全く問われなくなってしまう。また、「パフォーマンス評価」という語でもウィギンズにおいては不適当だろう。なぜなら、そうした場合、従来よく利用された多肢選択式からパフォーマンス型に形式が変わることを示すことはできても、彼の教育評価論の中で強調される文脈の現実性については問われなくなってしまうからである。このことは、彼の次のような発言に凝縮されていると言えよう。「『真正の』というのは『パフォーマンス評価』と同義ではない。いわゆるパフォーマンス課題や論述式テストの多くは、文脈内で大人が試される条件を模写していない。これが、『真正の』というような誤解される可能性のある言葉を使わなければならない理由である」(Newmann, Brandt, & Wiggins, 1998, p.21)。

このように、ウィギンズの「真正の評価」論は、評価の形式ではなく、それが行われる文脈を現実に近づけることに最たる眼目があったのである。

（2）文脈の妥当性

このような文脈へのこだわりは、彼の「妥当性」概念にも見られる。

文脈と妥当性に関連して、ウィギンズは、メシック（S. Messick）などの論考を援用しながら、次の点を強調する。「妥当性は、テストが行われる文脈――それは、そのテストが意図され

ていた目的のために使われていたかどうか、ということである――と、そこから生ずる結果に関して分析されなければならない」（Wiggins, 1993, p.239)。そこで、妥当性の検討に当たっては、「テストの解釈と使用における文脈の役割を現在の課題として繰り返し検討し監視する」（Messick, 1989, pp.14-15＝邦訳、p.23）というメシックの主張をウィギンズは重視する。つまり、文脈面から妥当性の検証を行うことに重点を置いたのであり、これこそ従来のテスト作成者に不問にされがちなところであった。

また、ウィギンズの「真正の評価」論においてその復活が謳われた（Wiggins, 1993, p.243）「表面的妥当性」は、テストが測定しているように思われるものを見るだけで、テストが実際に測定するものを正確に示すものではない。そのため、専門（技術）的な意味では「妥当性」とは言えない。精神測定学者（psychometrician）が表面的妥当性を嘲る理由もここにある。しかしながら、メシックも指摘するように、「テストが応答者、使用者、あるいは他の人の目からその目的に関係しているかどうかは、受験者の協力や動機づけ、あるいは使用者や一般大衆によるテスト結果の受け入れに影響を与える。したがって、表面的妥当性の欠如は可能な限り避けるべきである」（Messick, 1989, p.19＝邦訳、p.29)。「真正の評価」は大人が現実世界で直面するような難題の中で行われるものである。しかし、いくら文脈が現実に近づいていても、子どもの方に「現実世界に挑戦している」という感触がないと、その子どもはやろうとしない。表面的妥当性は、この感触の重要性を指摘するのに必要なのである。

このように、ウィギンズは妥当性の面からも文脈を重視していたのである。

（3）理解の文脈依存性

彼はなぜこれほどまでに文脈を重視するのだろうか。これを解く鍵は、ウィギンズの知識の捉え方から求めることができる。

知識と文脈との関係で、ウィギンズの次のような発言は傾聴に値しよう。「ある生徒が知識を『所有して（possess)』いるかいないか、どちらか一方で語ることは適切ではない。そうではなく、テストを受ける人が――文脈内で――聡明に、あるいは無知に行為するのである[8]」（Wiggins, 1993, p.241)。つまり、どのようなものであれ、そこで提出される「知識」が適当かどうかは全て文脈に依存するのである。ここに文脈を強調する論拠を求めることができるが、それは彼の「理解（understanding)」概念によって一層堅固なものとなる。

ウィギンズによれば、理解のテストとは、「知識を思慮深く適用することができるかどうかを見ること」であり、そこでは、知識の利用に関する「よい判断（good judgement)」が要求されるという。そして、「『よい判断者になるとは、ややこしい状況が持つ多様な特性の相対的な価値に関する勘を持つこと』(Dewey, 1933, p.123＝邦訳、p.126を改訳）であって、曖昧でない事例に何らかの形で明確に応用される一般原理に関する知識を単に持つことではない」（Wiggins, 1993, p.219)。換言するなら、「よい判断」、そして、それを通して明らかにされる「理解」は、何が正解とも分からない不明瞭な状況の中でしか見られないのである。これこそ、現実世界で起こっている問題状況でありながら、学校教育

において敬遠されてきたものである。ここに、現実的な文脈を求める最たる理由がある。

おわりに

以上、ウィギンズが提唱した「真正の評価」論の概要を示し、それに対する批判を踏まえて、「真正性」概念の特質を探究してきたわけだが、これらの作業から次のようなことが明らかになった。すなわち、ウィギンズの「真正の評価」論は、実世界の文脈を模写した現実的なシミュレーションの中で評価活動を行うことを基本とするもので、現実的な文脈が重視されるのは、「よい判断」を通して明らかにされる「理解」を見るためだった、ということである。

ウィギンズが「真正の評価」論を提唱したのは、単に標準テストやそれを擁護する精神測定学（psychometrics）を批判するためではなかった。彼が求めていたのは、あくまで、円熟した「精神の習慣」を母胎に「よい判断」を生み出す「思慮深い理解」だった。それを授業の中で実現するために、評価課題の文脈を現実に近づけておく必要があったのである。

最後に、本稿で取り上げたウィギンズの「真正の評価」論は、単にテスト論に終始するだけのものではなく、カリキュラム論と連動するものであった、ということを付け加えておきたい。

1980年代後半、ウィギンズは「思慮深さ」を目指した教育について具体的な構想を示しており、それを実現させるためのカリキュラム論を当時既に展開していた(Wiggins, 1987)。例えば、彼がその研究を主導していたエッセンシャル・スクール連盟（Coalition of Essential Schools）では、「本質的な問い」を軸にカリキュラムを編成することが試みられていたが（Cushman, 1989)、それこそ彼が構想していた「思慮深い教育」を具体化するものであり、それを最後まで貫徹させることに寄与した最終試験としての「学習発表（exhibition)」（Cushman, 1990）こそ、「真正の評価」のモデルとなっていたものである。したがって、ウィギンズの「真正の評価」論は、このような一連のカリキュラム改革の動きの中で生み出されたものであった、ということは確認しておかなければならないだろう。

では、当時ウィギンズは一体どのようなカリキュラム論を展開していたのだろうか。今後の課題としたい。

註

1）ウィギンズの教育評価論の体系については、Wiggins（1998）を参照のこと。ちなみに、日本では例えば田中耕治（2002）が、ウィギンズの所説を参照しながら「真正の評価」論を説明している。

2）当時（1989年)、ウィギンズは「真正のテスト」や「本当のテスト」という表現を使っているが、そこで想定されている「テスト」は、いわゆる「標準テスト」のことではない。「テスト」という言葉の語源について言及していることからも分かるように、彼は「テスト」という言葉をその原義、すなわち「人の努力の価値を決める全ての手続き」という意味で用いていた(Wiggins, 1989b, p.708)。

3）この課題は、ロードアイランド州プロヴィデンスにあるホープ高校（Hope High School）で使われていたもので、エッセンシャル・スクール連盟の「共通原理（Common Principles)」

の１つである「学習発表による卒業証書(Diploma by exhibition)」(Sizer, 1992, p. 226) に依拠したものである。ちなみに、エッセンシャル・スクール連盟においてウィギンズは、サイザー（T. R. Sizer）が提起した原理を実際のカリキュラムに具体化する役割を担っていた。なお、エッセンシャル・スクール連盟の取り組みについては安藤（1997）及び後藤(2002) を参照のこと。

4）ウィギンズによれば、「パフォーマンス」の動詞形「perform」は、「完成する (consummate)」あるいは「成し遂げる（accomplish)」という意味が原義にあることから、「課題を果たし、完成に持ち込む」という意味の言葉である。このことから、パフォーマンスは、知識や技能を用い、手近にある特定の文脈に対応しながら、自分自身の作品を生み出す中で検証されることになる（Wiggins, 1993, p.209)。

5）なお、第２点目は、教育評価における「衡平(equity)」をめぐる主張であるが、本節での対象とはしていない。この点については、Wiggins (1989b, pp.708-711) を参照のこと。

6）「モデレーション」とは、複数の教師が集まって、結果の比較や、評定規準の設定を行うことである。

7）なお、西岡加名恵らによるインタビューの中で、「ウィギンズは、『真正の評価』を行う際には、実際に現実の文脈の中で力を発揮することを求める場合と、シミュレーションされた文脈を作り出す場合があると述べている」(西岡、2003、p.142)。

8）これは、知的パフォーマンスを、「事実知(knowing that)」の「心的 (mental)」な応用ではなく、それ自体「技能知 (knowing how)」とする認識論によるものである（Ryle, 1949, pp.25-61＝邦訳、pp.23-78を参照)。

引用文献

安藤輝次、1997、「ポートフォリオ評価法によるカリキュラム改革と教師の力量形成（Ｉ）―エッセンシャル・スクール連合の試み―」『福井大学教育実践研究』第22号、pp.1-19。

Archbald, D. A. & Newmann, F. M., 1988, *Assessing Authentic Academic Achievemnt in the Secondary School*, National Association of Secondary School Principals.

Brown, J. S., Collins, A., & Duguid, P., 1989, Situated Cognition and the Culture of Learning, *Educational Researcher*, 18(1), pp.32-42＝杉本卓〔訳〕、1992、「状況に埋め込まれた認知と、学習の文化」石崎俊ほか〔編〕『認知科学ハンドブック』共立出版、pp.36-51。

Cizek, G. J., 1991a, Innovation or Enervation? Performance Assessment in Perspective, *Phi Delta Kappan*, 72(9), pp.695-699.

Cizek, G. J., 1991b, Confusion Effusion: A Rejoinder to Wiggins, *Phi Delta Kappan*, 73(2), pp.150-153.

Cumming, J. J. & Maxwell, G. S., 1999, Contextualising Authentic Assessment, *Assessment in Education*, 6(2), pp.177-194.

Cushman, K., 1989, Asking the Essential Questions: Curriculum Development, *Horace*, 5(5).

Cushman, K., 1990, Performance and

Exhibitions: The Demonstration of Mastery, *Horace*, 6(3).

Dewey, J., 1933, *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process* (New ed.), D. C. Heath＝植田清次〔訳〕、1955、『思考の方法―いかに我々は思考するか―』春秋社。

後藤武俊、2002、「米国エッセンシャル・スクール連盟の学校改革支援活動―『コミュニティとしての学校』理念を中心に―」『教育学研究』第69巻第2号、pp.205-214。

Messick, S., 1989, Validity, in Linn, R. (ed.), *Educational Measurement* (3rd ed.), American Council on Education／Macmillan＝池田央〔訳〕、1992、「妥当性」『教育測定学原著第3版 上巻』C.S.L.学習評価研究所（みくに出版）。

Newmann, F., Brandt, R., & Wiggins, G., 1998, An Exchange of View on "Semantics, Psychometrics, and Assessment Reform: A Closer Look at 'Authentic' Assessment", *Educational Researcher*, 27(6), pp.19-21.

西岡加名恵、2003、『教科と総合に活かすポートフォリオ評価法―新たな評価基準の創出に向けて―』図書文化。

Ryle, G., 1949, *The Concept of Mind*. University of Chicago Press＝坂本百大ほか〔訳〕、1987、『心の概念』みすず書房。

澤田稔、1997、「アメリカ合衆国における教育方法改革の最前線」松浦善満・西川信廣〔編〕『教育のパラダイム転換―教育の制度と理念を問い直す―』福村出版。

Sizer, T., 1992, *Horace's Compromise: The Dilemma of the American High School* (3rd ed.), Houghton Mifflin.

田中耕治、2002、「オーセンティック・アセスメントとは何か」『指導要録の改訂と学力問題―学力評価の直面する課題―』三学出版。

Terwilliger, J. S., 1997, Semantics, Psychometrics, and Assessment Reform: A Closer Look at "Authentic" Assessment, *Educational Researcher*, 26(8), pp.24-27.

Terwilliger, J. S., 1998, Rejoinder: Response to Wiggins and Newmann, *Educational Researcher*, 27(6), pp.22-23.

Wiggins, G., 1987, *Thoughtfulness as an Educational Aim*, Unpublished doctoral dissertation, Harvard University.

Wiggins, G., 1989a, A True Test: Toward More Authentic and Equitable Assessment, *Phi Delta Kappan*, 70(9), pp.703-713.

Wiggins, G., 1989b, Teaching to the (Authentic) Test, *Educational Leadership*, 46(7), pp.41-47.

Wiggins, G., 1991, A Response to Cizek, *Phi Delta Kappan*, 72(9), pp.700-703.

Wiggins, G., 1993, *Assessing Student Performance: Exploring the Purpose and Limits of Testing*, Jossey-Bass.

Wiggins, G., 1998, *Educative Assessment: Designing Assessments to Inform and Improve Performance*, Jossey-Bass.