

パフォーマンス評価による学習の質の評価 —学習評価の構図の分析にもとづいて—

松下 佳代

(京都大学高等教育研究開発推進センター)

Assessment of the Quality of Learning Through Performance Assessment: Based on the Analysis of Types of Learning Assessment

Kayo Matsushita

(Center for the Promotion of Excellence in Higher Education, Kyoto University)

Summary

The purpose of this paper is to analyze how the assessment of the quality of learning in higher education is conducted through performance assessment and to identify what conflicts arise between two paradigms.

First, to depict the structure of the arguments of learning assessment we set two axes, one of which was the axis of direct measures vs. indirect measures; the other was the axis of psychometrics paradigm vs. alternative assessment paradigm.

Next, we observed two trends in present performance assessment movements which reflect two contrastive paradigms, even though performance assessment was originally proposed under the alternative assessment paradigm. One trend is to construct an objective standardized test of performance assessment type, a representative example being the Collegiate Learning Assessment (CLA) adopted by the OECD's Assessment of Higher Education Learning Outcomes (AHELO). The other trend is to collaboratively develop performance assessments advanced by Alverno College and the Valid Assessment of Learning in Undergraduate Education (VALUE) Project of the American Association of Colleges & Universities (AAC&U).

Lastly, we produced the contrast between limited academic learning and the whole student engagement as the third axis of learning assessment arguments. We can analyze the types of present learning assessment activities along these three axes.

キーワード: パフォーマンス評価、学習評価、CLA、ルーブリック、VALUE プロジェクト

Keywords: performance assessment, learning assessment, CLA, rubric, VALUE Project

1. はじめに

(1) 学習評価への着目

わが国の高等教育政策の焦点は、近年、〈教員が何を教えるか〉から〈学生が何を学んだのか〉に明確にシフトしてきた。それを決定づけたのは、中央教育審議会の2つの答申、「学士課程教育の構築に向けて」(2008年12月)と「新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学～」(2012年8月)である。

〈学生が何を学んだか〉への焦点化は、実は今に始まったことではない。デューイ (Dewey, J.) は、すでに百年近くも前にこう述べている。「他の誰かが買うときに初めて売ることができるのと同じように、他の誰かが学ぶときに初めて教えることができるのだ」(Dewey, 1991, p. 214)。ここに表現されているのは、学生の学びに依存して、教員の教えるという行為が初めて成り立つという認識である(松下・田口, 2012)。

しかし、今日の学生の学習への焦点化がかつてと異なるのは、〈学生が何を学んだか〉を「学習成果 (learning

outcomes)」として評価することが、教育にとって不可欠の要素としてとらえられていることである。本稿では、学生の学習の「質」に注目し、その評価のあり方について検討する。

(2) 質とは何か

「質」の問題は、高等教育では、質保証の文脈で論じられることが多い。そこでの質は、関係者の立場（学生、教職員、政府、企業など）や高等教育のとらえ方によって意味の異なるきわめて多義的な概念であり（大学評価・学位授与機構，2011，p. 67）、質といいながら、実際は数量的に表現されることも多い。オランダの大学評価専門家フローインスティン（Vroeijenstijn, A. I.）は、「質は、質を求めるすべての関係者の間での交渉の産物」であり、高等教育機関はそうした要求を目的・目標に反映させるべきだとした上で、「目的・目標の達成」を質とみなしている（Vroeijenstijn, 1995, 邦訳 pp. 30-35）。また、質と卓越性（excellence）の違いについて、ファーストフードと星付きレストランを例に取りながら、「卓越性」が両者を同一の基準で比較した上でより優れていることを意味するのに対し、「質」は両者がそれぞれに異なった基準で「期待通りのものを得られるか」を追求することであると説明している。質と卓越性の区別をふまえるならば、「高等教育機関が、優秀な学生だけでなく多くの学生を教育するために、『卓越』を目指さないという選択をすることは可能である」（邦訳 p. 34）とフローインスティンはいう。つまり、質はすべての高等教育機関においてめざされるべきものだが、卓越性はそうでないということである。この質と卓越性の区別は、本稿のテーマとする学生の学習評価ではなく、大学評価に関するものだが、ユニバーサル段階にある各大学のミッションを考える上で重要な区別である。

一方、質的研究においては、質は「ある事物（モノや経験）が本来持っている特性や現象上の特性、あるいは本質的な特徴」（Schwandt, 2007, 邦訳 p. 105）と定義され、「質的」は「量的」と対照的な性格を表すことばとして用いられる。このような意味での質についての評価として「質的评价」という考え方がある。質的评价とは、「評価データ、評価基準、評価対象、評価結果の表現などにおいて、数量より質が重視される評価の総称」であり、「文字、絵・図、映像、音声、身体表現などの形式で質的データが収集・記述され、それらを用いて質的な評価基準の下で生徒の学習や教育活動などの質が評価され、さらに、その評価結果が文章などによって質的に表現される（数値によって表現されることもある）」（松下，印刷中）ような評価のことをいう。

(3) パフォーマンス評価

本稿の目的は、このような質的评价を中心に、学生の学習の質の評価について、その現状と論点を明らかにすることにある。その際、本稿では主な検討対象として「パフォーマンス評価（performance assessment）」を取り上げる。パフォーマンス評価とは、〈ある特定の文脈のもとで、さまざまな知識や技能などを用いながら行われる、学習者自身の作品や実演（パフォーマンス）を直接に評価する方法〉のことである。「生徒のパフォーマンスやパフォーマンスの事例を、設定されたパフォーマンスの基準に基づいて、直接かつ体系的に観察し評価すること」（Hart, 1994, 邦訳 p. 148）とも定義されている。

「パフォーマンス評価」という用語自体は目新しいものであったとしても、大学教育の中に、パフォーマンス評価に類するものは実は数多く存在する。たとえば、医療教育や教員養成などの分野での実技、芸術教育分野での演奏・演技や作品、PBL（Project Based Learning）での製作物、さらにいえば、レポート・論文や口頭発表などを通じて行われる評価は、どれもパフォーマンス評価になりうる。だが、これまでこうした評価は、多くの場合、個々の教員の主観にゆだねられていた。パフォーマンス評価の観点から検討することで、大学教育におけるこうした評価に評価論の光をあてることが可能になる。

日本では、パフォーマンス評価の研究は初等・中等教育の方が先行している（西岡・田中，2009；田中，2005，2011；松下，2007）。とりわけ、2008年の学習指導要領の改訂にともなう指導要録の改訂に際し、「思考力・判断力・表現力」の評価の方法としてパフォーマンス評価への言及がなされたことにより¹⁾、学校現場にも広がりを見せている。

一方、大学教育では、本来はパフォーマンス評価やポートフォリオ評価などにおけるツールの一つであるはずの「ルーブリック（rubric）」（複数の基準とレベル、それを説明する記述語からなる評価基準表）のみが一人歩きしている観が否めない。〈パフォーマンス評価＝ルーブリック評価〉という狭い理解に陥っている例もみられる。2012年8

月の中教審答申「新たな未来を築くための大学教育の質的転換に向けて」においてルーブリックが取り上げられたことにより²⁾、その傾向にいつそう拍車がかかるおそれもある。

（4）本稿の目的と構成

以上のような問題意識から、本稿では、パフォーマンス評価によって大学教育における学習の質の評価がどのような形で行われるのか、また、そこにどんな議論や対立があるのか、パフォーマンス評価は学習評価全体の構図の中にどう位置づけられるのか、を明らかにすることとする。

まず、大学教育における学習成果の評価の現状を概観し、学習評価の構図を描き出す。次に、パフォーマンス評価の特徴を述べ、それをふまえて、現在、大学教育において取り組まれているパフォーマンス評価に2つの方向性がみられることを指摘する。その一つは、パフォーマンス評価型の標準テストであり、もう一つは、組織的なパフォーマンス評価である。標準テストとしては、OECDのAHELO（Assessment of Higher Education Learning Outcomes：高等教育における学習成果の評価）プログラムにおけるジェネリック・スキルの評価で用いられているCLA（Collegiate Learning Assessment）を取り上げる。一方、組織的なパフォーマンス評価については、代表的な事例であるアルヴァーノ・カレッジ（Alverno College）でのパフォーマンス評価、および、大学間の共同開発の事例として、AAC&U（American Association of Colleges & Universities：アメリカ大学・カレッジ協会）のVALUE（Valid Assessment of Learning in Undergraduate Education：学士課程教育における妥当な学習評価）プロジェクトを取り上げることとする。最後に、これらの検討にもとづいて、学習評価の構図を再描出するとともに、パフォーマンス評価による学習の質の評価の課題について考察する。

2. 学習評価の構図

（1）直接評価と間接評価

日本の大学教育で、学習評価が広く注目されるようになったのは、2008年12月の中教審答申「学士課程教育の構築に向けて」において、学士課程修了段階での学習成果が「学士力」と名づけられ、その達成度の評価が求められるようになってからである。この答申は、日本の大学教育の世界に、〈教員が何を教えるか〉よりも〈学生が何を学んだのか〉に力点を置く「成果にもとづく教育（outcome-based education）」の考え方を公式に持ちこむことになった（松下，2012，pp.30-32）。

学習評価はまず、大きく「間接評価」と「直接評価」に分けることができる（山田，2012）。間接評価とは、学生の学習行動や自己認識を通じて——「どのように学習したか」や「何ができると思っているか」を学生自身に答えさせることによって——、学生の学習成果を間接的に評価することである。間接評価は、たいていの場合、集団を対象とする質問紙調査の形で行われる。北米では、インディアナ大学が開発したNSSE（National Survey of Student Engagement）や、UCLAが開発したCIRP（Cooperative Institutional Research Program）、CSS（College Senior Survey）などが知られている。日本でも、CIRPやCSSの日本版であるJCIRP、JCSS（山田，2009，2012）や、「全国大学生調査」（東京大学大学経営・政策研究センター，2008）、「大学生のキャリア意識調査」（溝上，2009）などが実施されている。最近では、こうした質問紙調査で得られるデータを、各大学が大学経営にとって必要なIR（institutional research）³⁾のデータとして利用するようになってきていることから、間接評価はいつそうの重みを与えられるようになっている。

一方、直接評価とは、学生の知識や行為を通じて——「何ができるか」を学生自身に提示させることで——、学生の学習成果を直接的に評価することである。直接評価は、プロジェクトから小テストにいたるまで大小さまざまな課題の評価や成績評価を通じて、すべての大学教員が日々の教育実践の中で行っている。直接評価に関する近年の改革の際だった傾向は、2007年の大学設置基準の改正において「成績評価の厳格化」に関わる条項⁴⁾が加えられて以降、GPA（Grade Point Average）制度が広く普及したことである。文部科学省（2011）によれば、2009年度現在、学部段階で360大学（約49%）、研究科段階で159大学（約27%）がGPA制度を導入している。その利用のしかたは、個別の学修指導への活用、奨学金等の選定基準や進級・卒業判定の基準としての活用などである。しかしながら、GPA自体は、成績（評価結果）の数値化・縮約化にすぎず、評価方法の改善に直接つながるものではない。本稿の主題で

あるパフォーマンス評価は、いうまでもなく直接評価の一方法である。パフォーマンス評価においては評価結果の表現だけでなく、評価課題や評価基準⁵⁾も議論の俎上にのせられ、教授・学習への影響はGPAよりはるかに大きいと考えられる。

(2) 評価の2つのパラダイム

さて、直接評価・間接評価という区別以上に、さまざまな評価の特徴を把握する上で重要なのが、評価のパラダイムの違いである。評価のパラダイムは、「心理測定学的パラダイム」とそれに代替・対抗・補完するものとしての「オルターナティヴ・アセスメントのパラダイム」に概念化することができる。オルターナティヴ・アセスメント（代替的評価）は、「標準テストを代替したり、補完したりするようにデザインされたさまざまな評価方法」の総称である（Hart, 1994, 邦訳 p. 145）。

① 背景

アメリカにおいて初等・中等教育でオルターナティヴ・アセスメントが主張されるようになった背景には、1980年代半ば以降、アカウンタビリティの重視により標準テストが多用されるようになったのに対して、それでは評価できない高次の（あるいは統合的な）能力を評価しようとする理論的・実践的動向があった。

大学教育についても、アメリカでは、標準テストの開発・実施の長い歴史がある。心理測定学の泰斗であるシェイベルソン（Shavelson, R. J.）は、その歴史を、標準テストの起源となる第Ⅰ期（1900～33年）、現在の一般教育の成果測定用の標準テストの基盤がつくられた第Ⅱ期（1933～47年）、テスト開発と測定に関わるテスト機関が増加した第Ⅲ期（1948～78年）、外部によるアカウンタビリティの要求が高まった第Ⅳ期（1979年～現在）に整理している（Shavelson, 2010；山田, 2012）。第Ⅲ期には、ETS（Educational Testing Service、1947年設立）のMAPP（Measure of Academic Proficiency and Progress）やACT（American College Testing、1959年設立）のCAAP（Collegiate Assessment of Academic Proficiency）などが開発され、第Ⅳ期には、シェイベルソンも共同開発者となり、CAE（Council for Aid to Education、1952年設立）によってCLA（Collegiate Learning Assessment）が開発されている。

一方、客観テスト以外の評価方法も大学教育では多く使われてきたが、前述したように、その評価はほとんどが教員個人の人々の主観にゆだねられていた。だが、質保証やアカウンタビリティの要請のもとで、より妥当性や信頼性の高い評価方法へと鍛え直していく必要性が生まれてきている。とりわけ、2006年に出された高等教育将来構想委員会（Commission on the Future of Higher Education）の答申（いわゆる「スプリングス・レポート」）において、連邦政府が高等教育機関に対し、学生が応分の学習成果を獲得していることを、より客観的かつ統一的な形で示すよう求める勧告が行われたことは、大きなインパクトを与えることになった（吉田, 2009；森, 2012）。

② 2つのパラダイムの特徴

心理測定学的パラダイムとオルターナティヴ・アセスメントのパラダイムの違いを対比的に整理すると、表1のようになる。

表1 評価の2つのパラダイム

	心理測定学的パラダイム	オルタナティブ・アセスメントの パラダイム
学問的基盤	心理測定学	構成主義、状況論、解釈学など
評価目的	アカウンタビリティ、質保証	教育改善・指導、学生の成長
評価対象	集団	個人
評価機能	総括的評価	形成的評価
評価項目	分割可能性	複合性
評価場面	脱文脈性 統制された条件	文脈性 シミュレーション、真正の文脈
評価基準	客観性	間主観性
評価データ	量的データ	質的データ
評価主体	評価専門家、政策担当者	実践者自身
評価方法	標準テスト、学生調査 など	真正の評価、ポートフォリオ評価、 パフォーマンス評価 など

2つのパラダイムの根底には、それぞれの学問的基盤の違いが横たわっている。心理測定学的パラダイムの基盤になっているのはいうまでもなく心理測定学であり、それは基本的には心理現象を数値化することによって成り立っている。一方、オルタナティブ・アセスメントの基盤はそれほど安定しているわけではないが、構成主義、状況論、解釈学などがその基盤を提供しているといえる（Gipps, 1994）。

次に評価目的・評価対象・評価機能についてみてみよう。「改善」と「アカウンタビリティ」は、評価全体を貫く2つの異質な理念である（Vroeijenstijn, 1995；松下, 2005）。たとえば、改善のための評価では、改善へのステップとして現状の問題点を見いだすことが望まれるが、アカウンタビリティのための評価では、できるだけ問題点は目立たぬようにして評価結果をよく見せようとする力が働きやすい。また、改善のためには個々の対象についての具体的な評価が有用だが、アカウンタビリティにおいては、対象となる集団全体についての総括的な縮減された評価が求められることが少なくない。

評価項目と評価場面における対比については、認知心理学者のレズニック（Resnick, L. B.）の指摘が想起される。レズニックは、心理測定学のテストには、知識は要素に分割できるという「分割可能性」と知の「脱文脈性」という2つの誤った仮定があると指摘し（Resnick, 1989）、それとは異なる評価方法としてパフォーマンス評価の開発に取り組んだ（Resnick & Resnick, 1996）。パフォーマンス評価で用いられるパフォーマンス課題では、それとは対照的に「複合性」や「文脈性」が重視されている（後述）。

評価基準についていえば、心理測定学的パラダイムでは、評価者の主観的判断が含まれない客観性が重視されるのに対し、オルタナティブ・アセスメントのパラダイムでは評価者の主観も排除されない。ただし、評価が恣意的になるのを防ぐために、主観のつき合わせや調整によって間主観性を担保することが求められる。

評価データに関しては、心理測定学的パラダイムでは当然のことながら量的データが用いられるのに対し、オルタナティブ・アセスメントのパラダイムでは、むしろ数値化されにくい学習成果をすくい取るために質的データが重視される。

心理測定学的パラダイムの代表的な評価方法は、多肢選択法や正誤法などを用いた客観テストタイプの「標準テスト（standardized test）」や、質問紙調査の形式で行われる「学生調査」である。オルタナティブ・アセスメントでは、「パフォーマンス評価」の他に、「ポートフォリオ評価（portfolio assessment）」や「真正の評価（authentic assessment）」が用いられる。また、前者の評価は、評価専門家によって開発・運用され、政策担当者によって利用されることが多いのに対し、後者の評価は、学生の教育に携わる実践者が評価主体になることが多い。

断っておくと、この2つのパラダイムはいわば理念型であって、現実に行われている評価は、両者の中間型であったり混在型であったりすることが少なくない。縦の列は互いに関連が深い、もう一方のパラダイムの項目と結びつ

く場合もある（たとえば、アカウントビリティを果たすのに質的データが使われる場合もあれば、教育改善・指導のために量的データが活用される場合もある）。さらにいえば、オルターナティブ・アセスメントのパラダイムは、心理測定学的パラダイムほど学問的に成熟しているわけでもない。

にもかかわらず、パラダイムをこのように整理するのは、オルターナティブ・アセスメントのパラダイムを心理測定学的パラダイムに対比される別のパラダイムとして概念化しなければ、心理測定学的パラダイムのみが評価論における唯一のパラダイムとしてみなされかねないからである。その場合、あらゆる評価は、心理測定学的パラダイムにのっとって、それにどれだけ忠実であるかによって価値づけられることになるだろう。逆にいえば、オルターナティブ・アセスメントをもう一つのパラダイムとして立てることで、心理測定学的パラダイムとは異なる性格をもった評価を正当に価値づけることが可能になる。両者の中間型や混在型という把握も、2つのパラダイムが概念化されていればこそ可能になるのである。

(3) オルターナティブ・アセスメントの評価方法

表1には、オルターナティブ・アセスメントの評価方法として、「パフォーマンス評価」の他に、「ポートフォリオ評価」と「真正の評価」をあげた⁶⁾。パフォーマンス評価が〈ある特定の文脈のもとで、さまざまな知識や技能などを用いながら行われる、学習者自身の作品や実演（パフォーマンス）を直接に評価する方法〉であるのに対し、ポートフォリオ評価は、〈ポートフォリオに収められた資料にもとづいて、学習者の成長のプロセスを評価する方法〉、真正の評価は、〈学習者に、仕事場や市民生活など現実世界の課題と類似した、本物らしさ（真正性）をもった課題に取り組ませる評価の方法〉と定義できる。

もっともこの3つの評価方法の区別は、概念的区別であって、主要には、どこに評価の焦点があるかの違いである。評価の焦点は、ポートフォリオ評価では、「ポートフォリオ」に収められた一定期間の成長プロセスを示す証拠資料に、真正の評価では、評価文脈や評価課題の「真正性」に、パフォーマンス評価では、ある文脈での「遂行」やその結果に置かれている。実際の評価場面では、この三者——とりわけ真正の評価とパフォーマンス評価——は重なり合うことも少なくない。たとえば、パフォーマンス評価が真正の評価課題や評価場面で行われるときは真正の評価になるし、一方、真正の評価はたいていの場合、作品や実演といったパフォーマンスの形をとる。また、ポートフォリオの中には、客観テストの答案や得点が含まれることもあるが、何らかの作品や実演の資料が含まれないポートフォリオは稀である。

本稿では、パフォーマンス評価を扱っているが、当事者によっては「パフォーマンス評価」とカテゴライズされていないものであっても、本稿の定義にてらしてパフォーマンス評価とカテゴライズできるものについては検討事例の中に含めている。

(4) 学習評価の構図

以上の議論をふまえて学習評価全体の構図を示すと、直接評価か間接評価か、心理測定学的パラダイムに立っているか、オルターナティブ・アセスメントのパラダイムに立っているかによって、図1のように大きく4タイプに分けることができる⁷⁾。

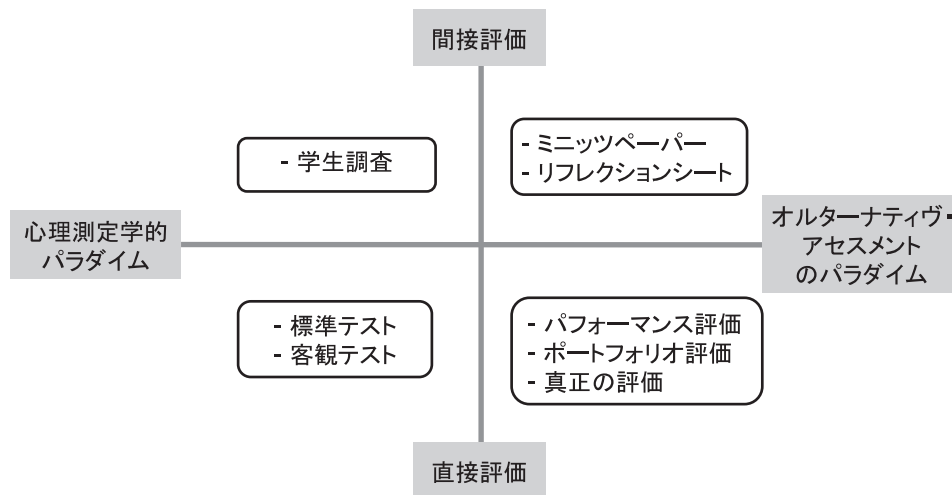


図1 学習評価の構図

本稿で扱うパフォーマンス評価は、オルタナティブ・アセスメントのパラダイムに立った直接評価としてひとまずは位置づけられる。とはいえ、近年では、CLAのようなパフォーマンス評価型の標準テストも用いられるようになっており、パフォーマンス評価内部でも入れ子のようにパラダイムの対立が生じている。以下では、そうした対立がどのような点で生じているのかを浮き彫りにしていこう。

3. パフォーマンス評価とは何か

(1) パフォーマンス評価の特徴

そもそもパフォーマンス評価とは、具体的にはどんな評価方法なのか。パフォーマンス評価とは何かを論じる上でまず必要なのは、パフォーマンスとは何かを論じることだろう。英語の“performance”には、「遂行」「成就」／「業績」「成果」／「演技」「演奏」／「できばえ」「性能」といった、多様な意味がある。「業績指標（performance indicator）」や「成果主義（performance-based pay system）」では、「業績」や「成果」と訳され、また、日常的には、「政治的パフォーマンス」のように、受けをねらった空虚な見せかけという意味あいでも使われている。

パフォーマンス評価でいう「パフォーマンス」は、「業績」や「成果」と無関係ではないが、空虚な見せかけという意味合いは本来含んでいない（しかし、そうなる危険性ははらんでいる）。実際に課題や活動を遂行させて、その遂行ぶり（実演）や遂行の成果物（作品）のできばえを評価するというのが、パフォーマンス評価の基本的な意味である。このような遂行の一形態である「演技」や「演奏」は、パフォーマンス評価一般を考える上でも参考になる。たとえば、フィギュアスケートの採点は教育以外の場面で行われるパフォーマンス評価の好例である（松下，2010a）。また、日常生活におけるさまざまな行為をパフォーマンス（演技）としてとらえたゴッフマン（Goffman, 1959）の理論も、パフォーマンス評価の意味や難点を考察する手がかりになる。

パフォーマンス評価には、①評価の直接性（パフォーマンスを実際に行わせて、それを直接、評価する）、②パフォーマンスの文脈性（パフォーマンスは具体的な状況の中で可視化され、解釈される）、③パフォーマンスの複合性（それ以上分割すると本来の質を失うという、一まとまりのパフォーマンスを行わせる）、④評価の分析性と間主観性（そうした質の評価のために評価基準と複数の専門家の鑑識眼を必要とする）といった特徴がある（松下，2010a）。

なかでも、評価課題である「パフォーマンス課題（performance task）」と評価基準である「ルーブリック（rubric）」は、パフォーマンス評価の特徴をよく表している。

(2) パフォーマンス課題

パフォーマンス課題とは、〈リアルな状況で、さまざまな知識や技能を総合して使いこなすことを求めるような課題〉のことである。パフォーマンス課題について、ウィギンズらは「効果的に行動するために知識を活用する課

題、あるいは、ある人の知識と熟達化を明らかにするような複雑な完成作品を実現する課題」(Wiggins & McTighe, 2005, 邦訳 p. 405)、ハートは「期待されている成果に[を]達成する能力を評価するために、生徒に与えられる課題」(Hart, 1994, 邦訳 p. 148)と定義している。これは狭義と広義の両端といえよう。ハートの場合は、「短いテストのようなものから、かなり長い時間をかけて展開される複雑なプロジェクトまで」のすべてをパフォーマンス課題と呼んでいるが、ウィギンズらは、短いテストというまでもなく、「批判的に考え、特定のアカデミックな応答、完成作品や実演を用意することを生徒に要求するような、オープンエンドの問いや問題」であっても、試験という条件下で内容が事前には公開されないような課題は、「アカデミック・プロンプト」と名づけて、「パフォーマンス課題」とは区別している。これでいえば、標準テストで課される「パフォーマンス課題」は「アカデミック・プロンプト」ということになる（逆に、自動車の路上試験は標準化されてはいるがパフォーマンス課題だといえる）。

ここでは、パフォーマンス課題には、具体的な状況の中で可視化され、解釈されるという「文脈性」や、それ以上分割すると本来の質を失うという一まとまりの「複合性」が求められるということをあらためて強調しておきたい。文脈性においてどの程度の真正性を求めるか、複合性においてどんな要素を含めるかについては幅があってよい。ウィギンズらは、パフォーマンス課題をデザインする際に考慮すべき要素として、「ゴール」「役割」「相手 (audience)」「状況」「完成作品・実演・意図」「成功を評価するスタンダードと規準」という6つの要素をあげている (Wiggins & McTighe, 2005, 邦訳 pp. 190f)。現実場面において人が何かを行うときには、何らかの問題や課題に対し、自分自身の意図をもち、何らかの役割を担いながら、誰か特定の相手に向けて、行為を行う。ウィギンズらのあるべき要素はそのような場面を構成する要素を模したものだといえるだろう。ゴッフマンは、日常生活における行為を劇場でのパフォーマンスとしてみることで、あらゆる行為が「自己呈示 (presentation of self)」の要素を含んでいることを暴き出したが、ウィギンズらは逆に、学校での評価場面に日常生活における自己呈示の要素を取り入れることでパフォーマンス課題を真正なものにしようとしているようにみえる。ただし、「成功を評価するスタンダードと規準」も示されているので、自分に対する他者の印象を望ましいものにしようとする「印象操作 (impression management)」——どうふるまえば高い評価が得られるかを考えながらパフォーマンスすること——が行われやすくなる。

こう考えると、このようなパフォーマンス課題は、現実世界の「真正性 (authenticity)」を模しながら、一方で、自分自身に忠実であろうとするという意味での「ほんものであること (authenticity)」(Taylor, 1991)を困難にするというジレンマに陥る危険性をはらんでいる。これは、パフォーマンス課題をデザインする際に留意しておくべきことだからである。

(3) ルーブリック

ルーブリックとは、パフォーマンス（作品や実演）の質を評価するために用いられる評価基準のことであり、一つ以上の基準（次元）とそれについての数値的な尺度⁸⁾、および、尺度の中身を説明する記述語からなる。基準 × 尺度のマトリックスで、各セルの中に記述語が入るという形式で表現されることが多い（具体的な事例については、巻末の資料1～3を参照）。

パフォーマンス評価で求められるパフォーマンスは文脈の中で表現される複合的なものである。評価基準の側に「分析性」が必要になる。そのような評価基準は、多くの場合、多次元的・多段階的である。それがルーブリックである。対照的に、客観テストでは、問題が細かく分割されているために、評価基準が正誤のような一次的・二値的なものです。また、客観テストは、その名の通り、誰が採点しても同一の結果が得られるという意味で客観的だが、パフォーマンス評価では、評価者の主観が入るので、評価者はパフォーマンスの質を適切に価値判断するための鑑識眼 (connoisseurship) を求められ、また、通常は複数の評価者がモデレーション (moderation) によって「間主観性」を担保することが必要になる。ここでいうモデレーションとは、複数の評価者間で評価の一貫性 (信頼性) を確保するために行う調整作業のことであり、評価過程を統一する方法と評価結果を統一する方法がある (Gipps, 1994; 松下, 2007, 2010a)。ルーブリックはそのようなモデレーションのためのツールにもなる。

もっとも、パフォーマンス評価にルーブリックが不可欠というわけではない。「分析性と間主観性」をみたま評価基準であれば、他の形式の評価基準でも、パフォーマンス評価は可能である。実際、自動車の路上試験や医療教育分

野でのパフォーマンス評価などでは、チェックリスト型の評価基準も用いられている（平山・松下，2009）。

ルーブリックはいくつかのタイプに分けられる。(a) 構造：基準を複数設定して分析的に評価する「分析的ルーブリック (analytic rubric)」か、基準を複数に分けずに全体的に評価する「全体的ルーブリック (holistic rubric)」か、(b) スcope：ある領域で一般的に適用できる「一般的ルーブリック (generic rubric)」か、当該課題だけに適用される「課題特殊のルーブリック (task-specific rubric)」か、(c) スpan：複数年にまたがって使われる「長期的ルーブリック (longitudinal rubric)」か、短期間あるいはスナップショット的に使われる採点用ルーブリックか、などである。

どのようなタイプのルーブリックを使うかは、パフォーマンス評価の性格を大きく左右する。これについては、次節以下で、具体的な事例を通してみていくことにしよう。

(4) 2つの方向性

2でもふれたように、現在では、パフォーマンス評価も、〈1〉標準化 (standardization) を志向するものと、〈2〉標準化を否定するもの、という2つの方向性に分化しつつある。標準化とは、すべての被評価者について運営上の条件と手順が統一されていることを意味する⁹⁾。2つの方向性のうち、一方は、パフォーマンス課題を大規模な標準テストの形式で与え、同一の評価基準で評価を行う。もう一方は、それぞれの大学・学科独自のパフォーマンス課題を開発し、枠組みを共有しつつも自分たちの大学・学科の固有性に応じて評価基準を作成している。前者が、パフォーマンス評価を心理測定学的パラダイムの下で洗練しようとするのに対し、後者はオルタナティブ・アセスメントのパラダイムを維持している。

なぜ、このようなことが生じるのか。ルーブリックは数値的な尺度を含むため、ルーブリックを用いることで、パフォーマンスの質は量的に表現されるようになる。いいかえれば、ルーブリックは質を量に変換する装置として機能する。つまり、ルーブリックそれ自体が、パフォーマンス評価をオルタナティブ・アセスメントのパラダイムと心理測定学的パラダイムに分化させる役割を果たしているのである。

本稿では、〈1〉の例として、アメリカのCAEの開発したCLAを取り上げ、〈2〉の例として、アルヴァーノ・カレッジにおけるパフォーマンス評価、およびAAC&UのVALUEルーブリックを取り上げる。前者は、パフォーマンス評価の標準化を図り、後者は、標準化に陥らないパフォーマンスの質の評価を追求している。

4. パフォーマンス評価型の標準テスト—CLA—

(1) OECD-AHELOにおけるジェネリック・スキルの評価

CLAのパフォーマンス評価が広く知られるようになったのは、「大学版PISA」といわれるOECD-AHELOのジェネリック・スキルの評価において、CLAが採用されてからだろう。

AHELOは、高等教育における学習成果の質を、国際的通用性と比較可能性をもった形で評価しようとする試みであり、2010～12年の間にその実行可能性を試すためのフィージビリティ・スタディが実施されている¹⁰⁾。「ジェネリック・スキル」「経済学」「工学」「付加価値 (value-added)」の4つの柱からなり、「付加価値」を除く3つがフィージビリティ・スタディの対象となっている。参加国は17カ国で（各国から約10大学が参加）、日本は「工学」に参加している。

「付加価値」は、高等教育に限らず、教育の質保証やアカウンタビリティの文脈において近年よく用いられるようになった概念である。AHELOは大学卒業時に実施されるが、入学時での大学間の学生の多様性を考えると、卒業時での測定だけでは教育の質を評価することができない。そこで大学在学中にどれだけ伸びたかによって教育の質を評価しようというのが、「付加価値」という考え方である。ただし、調査デザイン上の困難により、AHELOの現段階では、理論的・方法論的検討に限定されている¹¹⁾。「付加価値」という表現に、経済メタファーが用いられているのはいうまでもない（“quality assurance” がときには「品質保証」と訳されることもあるように、もともとと質保証という考え方そのものが経済メタファーの上で成り立っているのだが）。

ジェネリック・スキル、経済学、工学では、いずれもパフォーマンス課題が用いられている。ジェネリック・スキルでは90分で1問のパフォーマンス課題、経済学では1問のパフォーマンス課題（30分）と多肢選択問題（60分）、

工学では2問のパフォーマンス課題(60分)と多肢選択問題(30分)である。いずれのテストもウェブ上で実施される。

ジェネリック・スキルのテストでは、批判的思考、分析的推論、問題解決、文章コミュニケーションなどを統合的に用いながら、さまざまなソースからの資料を駆使して、現実的な状況の中のオープン・エンドな問いに答えることを求められる。

(2) CLA のパフォーマンス評価

①パフォーマンス課題

では、実際にどんなパフォーマンス課題が用いられるのだろうか。AHELO のフィージビリティ・スタディでどんな課題が用いられ、どのように採点が行われたのかはまだ公表されていない。ここでは AHELO 以外で実施されたことのある CLA のパフォーマンス課題の一つ、「犯罪削減 (Crime Reduction)」をみてみよう (Benjamin et al., 2009)¹²⁾。

学生にはまず、パフォーマンス課題に取り組むにあたっての注意(情報の正確性、考えの表現の明確さなど)によって評価されること、解答の分量に制限はないことなど)とともに、以下のような「ドキュメント・ライブラリ」が提示される。

ドキュメント 1: 与えられた指示

ドキュメント 2: 犯罪調査官のメモ

ドキュメント 3: 新聞記事

ドキュメント 4: 警察が公表した表

ドキュメント 5: 麻薬中毒者の治療プログラム XYZ に関する報告書

ドキュメント 6: 犯罪統計

ドキュメント 7: 対立候補イーガー博士の用いた図表

ドキュメント 8: 麻薬中毒者の治療プログラムに関する研究の要約

ドキュメント 1 には、次のようなシナリオと指示が書かれている。

バット・ストーンは、コロンビア州ジェファーソン市の市長選挙で再選をめざして運動中である。ストーン市長の対立候補はジェイミー・イーガー博士である。イーガー博士はジェファーソン市協議会の委員である。あなたはストーン市長の選挙コンサルタントである。

イーガー博士は、最近のテレビインタビューで次のような3つの主張を展開した。[中略。要約すると、①犯罪削減のために、市長のいうように警察官を増員したとしても、犯罪は増加するだけである、②警察官増員のための予算を麻薬中毒者の治療プログラムの予算に振り替えるべきだ、③麻薬中毒者の数を減らせば犯罪率も低下する、の3点]

ストーン市長は、イーガー博士のあげる3つの論点それぞれの長所と問題点を分析して、主張の欠陥を指摘したメモを準備するよう、あなたに依頼してきた。メモに求められるのは、イーガー博士のあげる3点に対するあなた自身の結論を含むこと、結論に対する根拠を説明していること、結論を出すにあたって依拠した特定のドキュメント・データ・叙述に言及しながらその結論を正当化していることである。(原文では、下線部は太字)

このシナリオに対し、3つの問いが与えられる。たとえば問1では、ドキュメント7にあるイーガー博士の用いた図表をみて、犯罪削減に関するストーン市長の提案は「犯罪を増加させるにすぎない」とするイーガー博士の主張の長所と問題点を述べつつ、博士の主張に賛成するかどうかを説明するよう求められる。

この CLA のパフォーマンス課題には、Wiggins & McTighe (2005) のあげるパフォーマンス課題の要素(ゴール／役割／相手 (Audience)／状況／完成作品、実演、意図／成功を評価するスタンダードと規準)のほとんどが盛り込まれていることが見てとれる¹³⁾。また、この課題では、PISA の読解リテラシーの問題にも似て、多様な種類のテキ

ストを読みこなし、批判的思考・分析的推論・問題解決を行い、自分の得た結論（＝問題解決の解）を文章で表現することが求められる。

もっとも、このシナリオは現実味があるとはいえ、シナリオに描かれた状況や役割を将来自分が遭遇するかもしれないものとして受けとめられる学生はごく一部だろう。そこに、多様性の進んだ大学教育の質を評価する課題としての限界もひそんでいる。

②採点基準と採点手続き

このようなパフォーマンス課題に対して、CLA では、4つの次元（分析的推論と評価、効果的な文章作成、文章作成の技法、問題解決）と6つのレベルからなる採点基準を設定している¹⁴⁾（巻末の資料1参照）。

CLA では、この採点基準を「ループリック」とは呼んでいないが、本稿の枠組みにてらしていえば、分析的で、一般的で、スナップショット的な採点用ループリックである。スナップショット的であるにもかかわらず一般的なループリックであることから、パフォーマンス課題の内容とは関係なく使われる、スキル重視のループリックだといえることができる。

採点は、2007年度は訓練を受けた採点者によって行われたが、2008年度からは機械採点も併用されている。さらに、CLA では、入学時と卒業時（希望すれば、中間点の2年次修了時も含められる）の変化をみることによって、付加価値の分析も行っている（Benjamin et al., 2009）。このように、CLA は、ジェネリック・スキルを総合的に評価するためのパフォーマンス課題と採点基準を提供しているだけでなく、ウェブ上での実施、採点の一部自動化、付加価値の分析などを備えている。AHELO に採用されたのもこうした点が認められたためだろう。

以上から、CLA において、パフォーマンス評価がどのような形で標準化されているかが明らかになった。この標準化の影響の及ぶ範囲は、CLA が AHELO のフィージビリティ・スタディで採用されたことによって、今後、OECD 加盟国全体やその他の国々にも広がる可能性がある。PISA が2009年には65の国・地域で実施され、世界経済の9割をしめる範囲までカバーするようになったのをみても、その可能性は小さくない。

では、このような標準化を志向するパフォーマンス評価に対し、標準化を否定するパフォーマンス評価はどのような形で展開されているのだろうか。CLA のような標準テストに明確に対抗することを意図して計画・実施されてきたのが、AAC&U の VALUE プロジェクトである。だが、その検討に入る前に、パフォーマンス評価の先進事例として知られるアルヴァーノ・カレッジの実践についてみておこう。

5. 組織的なパフォーマンス評価—アルヴァーノ・カレッジ—

(1)「能力をベースにしたカリキュラム」と「学習としての学生評価」

アルヴァーノ・カレッジは、ウィスコンシン州ミルウォーキーにあるカトリック系のリベラルアーツ・カレッジである（1887年創立¹⁵⁾）。学士課程は女子のみだが、共学の修士課程（経営学、看護学、教育学）も設置されている。学生数は大学院まで含めて2,605人（2011年秋学期現在）という小規模大学である。「能力をベースにしたカリキュラム（Ability-Based Curriculum）」や学習評価（パフォーマンス評価¹⁶⁾、ポートフォリオ評価、卒業生評価など）の取り組みで知られ（安藤，2006；川嶋，2008；松下，2010b）、VALUE プロジェクトでも12校のリーダーシップ校の一つとなっている（Rhodes, 2009b）。また、2-(1)でふれた NSSE の運営組織である NSSE Institute が実施している Project DEEP（Documenting Effective Educational Practice）¹⁷⁾でも、NSSE の調査結果から効果的な教育実践を行っている認められた Project DEEP Schools の一つに選ばれているほか、わが国の学士力や社会人基礎力の概念や取り組みにも影響を与えている（中央教育審議会，2008；経済産業省，2008）。

興味深いことに、アルヴァーノは、A～Fの成績評定（letter grade）やそれを数値化し平均した GPA を用いていない。評価は、授業科目ごとに、一定の基準（後述）にてらして可・不可（satisfactory と unsatisfactory）で判断され、学生へのフィードバックに力が注がれている。また、学生の実演や作品は e ポートフォリオ（Diagnostic Digital Portfolio：DDP）上に記録され、それがパフォーマンス評価の対象になるとともに、学士課程4年間にわたるその記録は、大学院進学や就職の際に、自分が大学で何を学んだかを説明するための証拠資料（evidence）として用いられる。このように、アルヴァーノでは、パフォーマンス評価やポートフォリオ評価において、量的方法より質的方法が選ばれ、その評価結果や評価資料が、教育改善や指導のためだけでなく、アカウンタビリティの遂行のためにも用い

られている。

我々にとっての評価とは、[中略] 基準にもとづいて、改善とアカウンタビリティのために行われる、学生の学習成果の評価のことである。[中略]

学習としての学生評価 (student assessment-as-learning) は、学習プロセスの中に埋め込まれており、また主に、学生の学習の改善に向けられているという意味において、「形成的」である。しかしながら、ここでは、「形成的」という語の意味を、アカウンタビリティも内包できるように拡張している。個々別々のハイ・ステイクスな評価としてではなくむしろ累積的なプロセスとして、その評価は個々の学生にとって「総括的」にもなるのである (Loacker & Rogers, 2005, p. 4)。

こうした評価の基盤となっているのが、「能力をベースにしたカリキュラム」である。1973 年にはじめて導入され、以後少しずつ改訂されながら現在に至っている。そのフレームワークは、機関全体を通して身につけさせる 8 つの能力と 6 つのレベルからなる (巻末の資料 2 参照)。8 つの能力とは、「コミュニケーション」「分析」「問題解決」「意思決定における価値判断 (Valuing in decision making)」「社会的インタラクション」「グローバルな視野の発達 (Developing a global perspective)」「効果的な市民参加 (Effective citizenship)」「美的な関わり (Aesthetic engagement)」である (Loacker & Rogers, 2005)¹⁸⁾。「コミュニケーション」「分析」「問題解決」「社会的インタラクション」はジェネリック・スキルのリストにしばしば含まれる能力だが、「意思決定における価値判断」「グローバルな視野の発達」「効果的な市民参加」「美的な関わり」には、カトリック系リベラルアーツ・カレッジとしてのアルヴァーノの特色が感じられる。

6 つのレベルは大きく初級レベル、中級レベル、上級レベルに分けられ、上級レベルでは、専攻する学問分野や自分の選んだ専門職との関係が密接になる。学士課程 4 年間で、8 つの能力すべてにおいてレベル 4 まで達することが卒業要件である。

この 8 つの能力と 6 つのレベルは、機関のミッションに即した長期的ルーブリックとして機能しており (ただし、「ルーブリック」という表現は用いられていない)、カリキュラムデザインにおけるフレームワークであるだけでなく、学生にとっての学習目標としても機能することが期待されている。

(2) 評価のレベルと評価組織

このルーブリックは機関 (大学) 全体で用いられる一般的ルーブリックであり、資料 2 にみるとおり抽象度も高い。したがって、これを各専攻において、さらには各授業科目において具体化していく必要がある。

我々は、これら 8 つの能力を、学生が個別特殊な文脈において発達させ、評価される必要があるということを認識しているが、まずは 8 つの能力をできるだけ一般的に記述し、それを各科目の担当教員や学科が、自分たちの学問分野の言葉で再記述できるようにした。(Alverno College Faculty, 1994, p. 9)

図 2 は、機関から専攻 (例: 歴史学科)、科目 (例: 「古代世界史」) へと具体化されるプロセスを示したものである。このように 8 つの能力の中から、その専攻に関連する能力が選ばれ、専攻の内容にあわせて具体化され、さらにその中から個々の科目に関連する目標 (期待される学習成果) が選ばれ、科目の内容にあわせて具体化される。

各科目では、その目標に即して評価課題と評価基準が設定される。たとえば、「古代世界史」(3 年後期) では、「何らかの基本的仮定の歴史的起源を説明するために歴史家が用いた概念・理論を選択し、その枠組みを別の時代や地域に適用して枠組みの説明力を検証しなさい」という内容の評価課題が出されている。形式とオーディエンスは、自分の知見をクラスの仲間に対して説明する口頭発表 (ビデオ録画)、仲間によるピアレビュー、研究レポート、解釈を行うために用いた仮定とプロセスについての自己評価報告である。

このような学生のパフォーマンス課題を、教員は「1. 問いを掲げ、論証する論題を提案するために、説明的枠組みを展開している」、「2. 記述的な証拠を他の証拠や枠組みと結びつけながら、理路整然とした議論を行っている」、

「3. 自分の見方が、情報源の選択や解釈にどう影響を与えているかを明確化している」、「4. オーディエンス（聞き手）のバックグラウンドや経験にあわせて、アイデアや情報をたえず調整しながら、語りかけている」、「5. 批判的コミュニケーションやアイデアの評価を行いやすくするために、一貫した書式（シカゴマニュアル）を用いている」という5つの基準によって評価する（Loacker & Rogers, 2005, p. 8）。この科目では、コミュニケーション、分析、問題解決、意思決定における価値判断、社会的インタラクションといった能力が、古代世界史という文脈の中で形成され評価されていることがわかる。

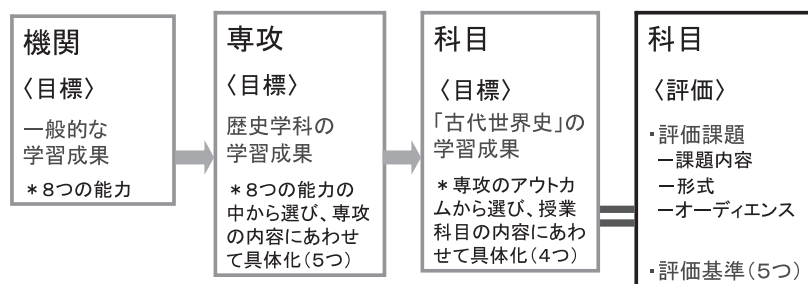


図2 機関・専攻・科目の目標と科目での評価の関係
（出典：Loacker & Rogers, 2005, p. 8 をもとに図式化）

では、こうした評価はどのような組織体制に支えられているのだろうか¹⁹⁾。アルヴァーノでは、ほとんどの教員は、学科（discipline department）と能力部門（ability department）の両方に所属することになっている。各能力部門は、毎年、年間目標を立てて活動を行っている。たとえば、ある年の「効果的な市民参加」部門の目標は、「新しく改訂されたレベル1・2に妥当するための条件を明確化し、それを達成するための授業科目を特定すること」であった（Alverno College Faculty, 1994, p. 91）。「学生評価協議会（Council for Student Assessment）」は、学科や能力部門と協働して、「学習としての学生評価」の企画・実施・運営の監督を行う（Loacker & Rogers, 2005, p. 3）。「教育研究・評価部（Educational Research and Evaluation department：ERE）」は、研究機関の性格をもった組織であり、アルヴァーノの教育・評価（学生評価、プログラム評価、機関評価、卒業生評価など）の枠組みの構築やレビューなどを行ってきた。さらに、「研究・評価協議会（Research and Evaluation Council）」は、管理職教員、学務部のアドミニストレーター、ERE メンバーからなる、研究・評価に係る最上位の組織であり、1976 年以来、ERE の長としてアルヴァーノの教育・評価研究を先導してきたメンコウスキー（Mentkowsky, M.）が、議長をつとめている。

一方、学生・教員に対する評価支援のサービスセンターとして機能しているのが、「評価センター（Assessment Center）」である。評価センターは、多様な機能をもった組織である。アルヴァーノでは、学外の専門家に、学生の地域やインターンシップ先での活動の評価を依頼しているが、そうした学外の評価を管理・運営し、教員を支援することも、評価センターの仕事の一つである。学生に対しては、入学前に「コミュニケーション」のレベル1の評価を行うことから始まり、学生のeポートフォリオ（DDP）の作成や自己評価を支援する。学生のパフォーマンス（作品や実演）の記録はすべて評価センターに保管されており、学生はいつでもアクセスできるようになっている。

（3）オルタナティブ・アセスメントのパラダイムの下での組織的パフォーマンス評価

以上みてきたように、アルヴァーノ・カレッジでは、1970 年代前半から 40 年近い時間をかけて、8つの能力と6つのレベルをフレームワークとする「能力をベースにしたカリキュラム」が構築され、パフォーマンス評価やポートフォリオ評価が実践されてきた。8つの能力と6つのレベルは、機関全体の一般的な長期的ルーブリックとして機能しており、各専攻や各科目での評価基準の土台ともなっている。

アルヴァーノでの組織的な評価は、今日の多くの大学のように、質保証やアカウンタビリティの要請に対応するために始まったのではなく、学生が自分の能力を把握し学習を進めていく手がかりとして導入された。「学習としての学生評価」は、こうした評価観を最もよく表す語である。そこでは、「個々の学生のパフォーマンスを、明示的な基準にもとづいて観察・判断し、それを学生にフィードバックすることを含む、学習にとって不可欠なプロセス」

(Alverno College Faculty, 1994, p. 3) として評価がとらえられている。

評価結果の表現では、成績評定や GPA は用いられず、長期的ルーブリックも、質を量に変換する装置として用いられてはいない。むしろ、アルヴァーノでは、アルファベットや数字に縮減されない形でパフォーマンス（作品や実演）そのものを、学生の成長の記録として e ポートフォリオ上に蓄積していくことが重視されている。実際に、こうしたパフォーマンスは、親や進学先（大学院）や就職先に対し、数値以上に学生の成長ぶりや大学の教育力を示すものとして、アカウンタビリティを発揮している。

このように、アルヴァーノの学習評価は、アカウンタビリティの要請にも応えられるよう拡張を図りつつ、明確にオルターナティブ・アセスメントのパラダイムに立つものだと見える。

6. ルーブリックの組織的開発—AAC&U の VALUE プロジェクト—

(1) VALUE プロジェクトの概要と目的

アルヴァーノ・カレッジの実践などを土台に、AAC&U が 2007 年から 2010 年まで取り組んだのが、VALUE (Valid Assessment of Learning in Undergraduate Education: 学士課程教育における妥当な学習評価) プロジェクトである²⁰⁾。AAC&U は、1915 年に設立された、教養教育の質や公共的役割の向上を目的とする組織であり、コミュニティ・カレッジから名門私立大学まで、タイプや規模の異なる 1,250 あまりの加盟校によって構成されている²¹⁾。

VALUE プロジェクトは、教養教育に関するルーブリックを共同開発し（＝「VALUE ルーブリック (VALUE rubrics)」）、e ポートフォリオと組み合わせることによって、学習の質の評価を行おうというものである。プロジェクト長であるローズ (Rhodes, T.) は、その目的を次のように述べている。

VALUE ルーブリックは、大学や教員が、特定の学位レベル——準学士号と学士号——の達成に対して期待されるパフォーマンスの質を評価するために、共通のスタンダードを創り出す手段を提供している。つまり、そこにあるのは、標準化に陥らない、質についてのスタンダードの枠組み (framework of quality standards without standardization) なのである。(Rhodes, 2011, p. 4)

ここには、「標準テストに対する妥当なオルターナティブ (valid alternative to standardized testing)」²²⁾を開発し広めようという VALUE プロジェクトの目的がはっきりと読みとれる。

2-(2) でふれたように、2006 年の「スプリングス・レポート」によって、アメリカの高等教育機関に対し、学生が応分の学習成果を獲得していることを、より客観的・統一的に示すことが求められるようになり、アカウンタビリティへの要請がいつそう高まることになった。従来、大学の教育力は、リテンション率（学業継続率）、卒業率などで示されることが多かったが、それらは、学生の学習成果の間接的指標にすぎず、アカウンタビリティの要請に十分応えることができない。そこで、直接的指標としてよく用いられるようになったのが、ACT の CAAP、ETS の MAAP、そして CAE の CLA といった標準テストの得点である。VALUE プロジェクトは、そうした状況のなかで、標準テストに対抗する学習評価の代替的アプローチとして構想されたものであった（森，2012）。

Rhodes (2009a) は、標準テストの問題点として、サンプルになった学生のみが対象であること、スナップショット的な学習成果の把握のしかたであること、評価結果が教授・学習の質の改善に役立たないことをあげる。逆にいえば、VALUE プロジェクトにおける学習評価は、当該機関のすべての学生を対象とし、長期的な学習成果を把握し、評価結果が教授・学習の質の改善に役立つようなものでなければならない。そのような学習成果の評価方法として選択されたのが、e ポートフォリオに収められた学生のワーク（活動や作品）を VALUE ルーブリックで評価するという方法であった。

e ポートフォリオが選ばれたのは、すでに多くの大学で広く利用されており、共有が容易であり、多様な表現が可能であることによる (Rhodes, 2009b)。VALUE プロジェクトでは、正課のカリキュラムだけでなく正課外の補助カリキュラム (cocurriculum)²³⁾も含む大学生活全体におよぶ学生のワークこそが、学生の学習を最もよく表す妥当な評価データとみなされている。そのようなワークを収めるには e ポートフォリオがよい媒体になると判断されたのである。VALUE プロジェクトでは、補助カリキュラムでの学習成果にも評価対象が広がられているが、これは、標準

テストの評価対象の狭さを批判的に克服しようとするものであると同時に、NSSE にみられるように、大学生活全体への学生の関わり（student engagement）を重視する大学教育の動向を反映したものであるといえよう。

（2）プロジェクトの組織と開発プロセス

VALUE プロジェクトは、VALUE リーダーシップ校 12 校（アルヴァーノ・カレッジ、ポートランド州立大学、スヘルマン・カレッジ、ミシガン大学など）、および VALUE パートナー校 70 校（オーバーリン・カレッジ、ジョージタウン大学など）によって開発された。リーダーシップ校は、e ポートフォリオを使用してきた大学であり、VALUE ルーブリックをテストし、ルーブリックの有用性・問題点・長所をフィードバックする上で中心的な役割を担った。一方、パートナー校の多くは e ポートフォリオこそ使っていなかったが、別の形で学生のワークを収集していた大学であり、リーダーシップ校と同じく、ルーブリックのテストに参加した。

開発は 3 つのラウンドで行われた。第 1 ラウンドでは、2008 年春に 3 つのルーブリックが作成され、リーダーシップ校でパイロットテストが実施された。第 2 ラウンドでは、2008 年秋に 14 のルーブリックが作成され、リーダーシップ校とパートナー校でパイロットテストが実施された。さらに第 3 ラウンドでは、改訂されたルーブリックについてパイロットテストが行われ、その結果を受けて、2009 年に全部で 15 の VALUE ルーブリックが公開されるに至った（Rhodes, 2009b）。

（3）VALUE ルーブリックの特徴

VALUE ルーブリックの開発された 15 の領域とは、①探究と分析、②批判的思考、③創造的思考、④文章コミュニケーション、⑤口頭コミュニケーション、⑥読解、⑦量的リテラシー、⑧情報リテラシー、⑨チームワーク、⑩問題解決、⑪市民参加、⑫異文化知識・能力、⑬倫理的推論、⑭生涯学習の基礎とスキル、⑮統合的学習である。①～⑩は知的・実務的スキル、⑪～⑭は個人的・社会的責任、⑮は単独で統合的・応用的学習、に大別されている（Rhodes, 2010）。これらは、教養教育によって形成されるべき能力として抽出されたものである。15 領域の VALUE ルーブリックについては、巻末の資料 3-1～3-15 を参照していただきたい²⁴⁾。

VALUE ルーブリックは、アルヴァーノの〈8つの能力と6つのレベル〉と同じく、一般的な長期的ルーブリックである。機関レベルで学生の学習を評価し議論することを目的としたものであり、CLA のように採点を目的としたものではない。各領域の能力は、ベンチマーク（1）、マイルストーン（2、3）、キャップストーン（4）の4つのレベルに尺度化されており、各数字は対応する学年を大まかに表している。1年生と4年生の得点の分布を比較することによって、標準テストとは異なる形で「付加価値」を評価できるようデザインされているわけである。一方、アルヴァーノのルーブリックが個々の能力については全体的ルーブリックであるのに対し、VALUE ルーブリックは、複数の基準（次元）を含む分析的ルーブリックとなっている（たとえば、両者共通に含まれている「問題解決」で比較すればその違いは明らかだろう）。

だが、何よりも注目すべき VALUE ルーブリックの特徴は、このルーブリックがメタルーブリックとしての性格をもっているということである（図3参照）。各大学で開発されてきた既存のさまざまなルーブリックをもとに、AAC&U の専門家チームによって VALUE ルーブリックが共同開発され、それが個々の大学・学科・科目の文脈にあわせてローカライズされる。その意味で、VALUE ルーブリックは、CLA のような採点用ルーブリックではなく、アルヴァーノのような機関全体のルーブリックでもなく、もう一段階抽象度の高い「メタルーブリック」である。あるいは、ルーブリックの元になるルーブリックという意味で「プロトルーブリック」と呼んでもよいかもしれない。VALUE ルーブリックは、これによって、大学をこえた「共通性」と大学ごとの「多様性」の統一を図っているといえる。本稿の冒頭で、フローインスティンによってなされた質と卓越性の区別について述べた。CLA のような標準テストが、各大学の教育力を同一の基準で比較することにより結果的に卓越性を競わせることになるのに対し、VALUE プロジェクトは各大学（群）による質の追求を促すものであるということもできよう。

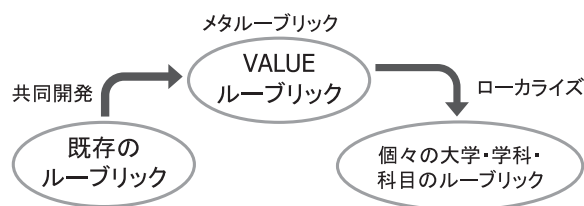


図3 VALUE ルーブリックの位置づけ

AAC&U では現在、VALUE プロジェクトの拡張として、CAAL（Collaborative for Authentic Assessment and Learning：真正の評価と学習のための協働）プログラムが始まっている。このプログラムは名称にも表れているとおり、真正の学習を促すような真正の評価をめざすものであり、標準テストへの批判もいっそう明確になっている。

現代の教育文化において、標準テストは往々にして、学生の能力や知識を測定するための、信頼性のある唯一の測定尺度と考えられている。しかしながら、多くの研究が示すように、標準テストそれ自体は、深い学習（deep learning）を促さないし、また、最も価値ある教育成果（問題解決やチームワークなど）を測定することもない²⁵⁾。

VALUE プロジェクトでもすでに、現実場面における学生のワークをポートフォリオ化することは含まれていたが、CAAL ではあらたに、大学・学科等の多様性をこえて、学生のパフォーマンスについてのベンチマーク・データ（標準となるデータ）を構築することがめざされている。具体的には、妥当化されたルーブリックにそって評価された学生の作品をその得点データとあわせて収集し、リポジトリを作成するという計画である。

これは、パフォーマンス評価でいう「アンカー（anchor）」の収集にあたる。アンカーとは、ルーブリックの各基準・レベルに対応するパフォーマンスの実例のことであり、採点の信頼性を高める上で不可欠なものとされる（Wiggins & McTighe, 2005, 邦訳 p. 389 参照）。Gipps（1994）は、心理測定学的パラダイムの下での「信頼性」の概念を、オルターナティブ・アセスメントのパラダイムの下では、複数の評価者間で評価の一貫性が確保されていることを意味する「比較可能性（comparability）」という概念に置き換えている（邦訳 p. 241）²⁶⁾。CAAL は比較可能性を高めることによって、信頼性に対する要求に応えようとするものだとみることができる。

このように、VALUE プロジェクトは、VALUE ルーブリックの開発からさらに CAAL プログラムへと拡張することによって、妥当性の高いオルターナティブ・アセスメントの構築への歩みを進めている。

7. 結び—学習評価のダイナミクス—

(1) 学習評価の構図—パラダイムの対立と相互浸透—

本稿で論じてきたことにもとづいて学習評価の構図をあらためて描くと、図4のようになる。

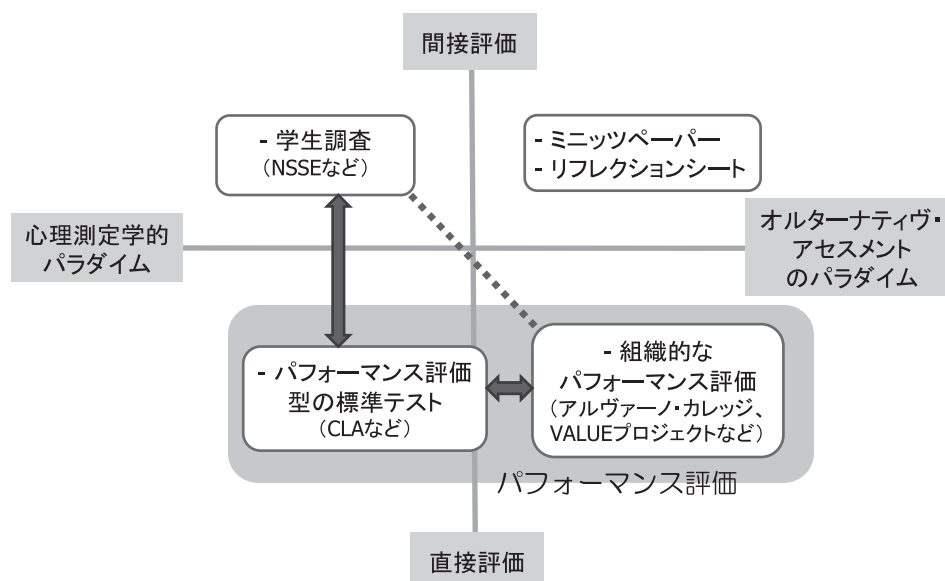


図4 パフォーマンス評価内部でのパラダイムの対立

①パラダイムの対立

CLAなどの標準テストにおけるパフォーマンス評価と、アルヴァーノ・カレッジ、VALUEプロジェクトなどにおけるパフォーマンス評価の間には、明確なパラダイムの対立がある。両者の対立は、とりわけ標準化を志向するかどうかにおいて際だっている。本論の中では、VALUEプロジェクトでの標準化批判を紹介したが、一方、CLAの方も、ポートフォリオによる真正の評価に対して、「ポートフォリオは、一見したところでは、学生のワークを提示するための『真正の』手段であるが、測定道具がもつべき標準化という決定的に重要な特徴を欠いている」(Benjamin et al., p. 4)と批判している。

CLAが、すべての大学の学生に対して同一の課題を課し同一の採点基準（ルーブリック）を用いるという点で標準化されているのに対し、VALUEプロジェクトの場合は、VALUEルーブリックやCAALによって枠組の共通性や採点結果の比較可能性を高めながらも、大学による多様性を積極的に認めている。各大学の内部でも、アルヴァーノ・カレッジの例でみたように、学科・専攻や科目によって評価基準は多様である。

また、評価内容についてみても、CLAの場合は、90分で解答させる限られた数の論述課題であるのに対し、アルヴァーノやVALUEプロジェクトでは長期にわたって蓄積された学生のパフォーマンスやワークであり、標準化より真正性が重視されている。

さらに評価結果の表現に関しても、パフォーマンスに対しルーブリックを用いて数値化することよりむしろ、一定の基準を満たすと判断されてポートフォリオに収められたパフォーマンスそのものを学生の学習成果の質の表現とみなしている点に、オルタナティブ・アセスメントとしての特徴が認められる。

②相互浸透

ただし、両者の間には、対立だけでなく相互浸透も見とれることできる。たとえば、CLAは現在、スナップショット的な標準テストから、教室実践での使用をめざすCLA in the Classroomへの拡張を図っている。

CLAの主な目標は、教員による教授・学習の改善を支援することにある。評価道具を支持（あるいは批判）する議論は、信頼性や妥当性といった心理測定的な問題やアカウンタビリティについての政策論争をめぐる展開されることが多く、したがって、評価道具が教室の中で教員にとって直接的に有用かという問いは、せいぜいのところ周辺的なものにとどまっていた。ここでは、CLAが教授・学習の改善にどのくらい使われているか（あるいは、使えるか）を示すために、このようなCLAのレリバンスについての問いを中心にすえることにしたい。(Benjamin et al., 2009, p. 1)

他方、アルヴァーノや VALUE プロジェクトでも、アカウンタビリティの要請にどう応えるかは重要な課題として認識されており、機関全体の長期的ループリックが学生の学習成果と大学の教育成果を実証する際のツールとして位置づけられている。

③第三の軸

ところで、図4でみると、NSSEなどの学生調査と、アルヴァーノ・カレッジや VALUE プロジェクトなどの組織的なパフォーマンス評価（とくにパフォーマンスのポートフォリオ化）とは、両軸で対照的な位置にあるが、第三の軸を立てると、両者の共通性もみえてくる。それは、主に正課のカリキュラムでのアカデミックな学習成果のみを評価するのか、それとも、正課・正課外の両方を含む学習成果全体、いかえれば大学生活への学生の関わり（student engagement）を評価するのか、という軸である。この軸では、CLAが前者であるのに対し、NSSEなどの学生調査とアルヴァーノ・カレッジや VALUE プロジェクトなどのパフォーマンスのポートフォリオ化は後者に位置づけられる。

カリフォルニア大学バークレー校 CSHE（Center for Studies in Higher Education：高等教育研究センター）のダグラスら（Douglass et al., 2012）は、彼らが大規模研究大学向けに開発した、大学生活への学生の関わりについての学生調査（SERU-S）と CLA を比較しながら、「CLA は『学習成果レース』に勝利し、米国内でも世界的にも『ゴールド・スタンダード』になりつつある」（p. 318）が、「CLA が、高等教育における妥当なアカウンタビリティにとって重要なのは、大学を得点化しランク化することを可能にする『付加価値』の統計的テストだと主張すればするほど、CLA はより広い意味での『システムの妥当性』を失うことになるだろう」（p. 324）と批判する。システムの妥当性とは、教育システムの目標を支えるのに役立つ情報や建設的な対応を生み出すという意味での妥当性である。

(2) 今後の課題

以上みてきたように、現在、学習評価の中では、評価の直接性、評価のパラダイム、評価する学習成果の範囲という3つの軸をめぐり、それぞれの軸で棲み分け・対立・相互浸透が生じている。なかでも、パフォーマンス評価の内部で評価のパラダイムをめぐってどのような対立と相互浸透がみられるかを、本稿では明らかにしてきた。

もっとも、本稿では、学習評価のダイナミクスを分析するにとどまり、これをふまえてどのような学習評価を構築すべきかについては論じることができなかった。教育評価が満たすべき要件として、本稿で言及した妥当性や信頼性（あるいは比較可能性）の他に、公平性（equity）や実行可能性（feasibility）があげられる（Gipps, 1994；Wiggins & McTighe, 2005；西岡, 2005）。これは、そのまま学習評価にもあてはまる。アルヴァーノ・カレッジや VALUE プロジェクトでとられているパフォーマンスのポートフォリオ化とループリックによる評価という方法は、高次の能力について広い範囲をカバーし、評価場面と実際に能力の発揮が求められる場面が近いという点で妥当性（内容的妥当性、表面的妥当性）は高いが、信頼性（あるいは比較可能性）、公平性、実行可能性という点では、なお難しさを抱えている。標準テストがこれだけ批判されながらも幅広く採用されているのは、これらの要件を満たしやすいためだろう。オルタナティブ・アセスメントとしてのパフォーマンス評価という形で学習の質の評価を追求するのであれば、信頼性（比較可能性）、公平性、実行可能性の要求にも応えていく必要がある。

私たちはすでに、理学療法教育（平山・松下, 2009；平山他, 2012）や歯学教育といった専門分野での指導、初年次教育におけるライティング指導（小野, 2012）などで、真正のパフォーマンス評価の開発を協働的に進めている。本稿はそうした開発のための理論的土台の構築と課題の明確化という意味をもつものである。

註

- 1) 中央教育審議会教育課程部会「児童生徒の学習評価の在り方について（報告）」（2010年3月24日）において、「思考力・判断力・表現力等を評価するに当たって、『パフォーマンス評価』に取り組んでいる例も見られる。パフォーマンス評価とは、様々な学習活動の部分的な評価や実技の評価をするという単純なものから、レポートの作成や口頭発表等による評価するという複雑なものまでを意味している。または、それら筆記と実演を組み合わせたプロジェクトを通じて評価を行うことを指す場合もある」と言及されている。
- 2) 答申の「用語集」では、ループリックを次のように説明している。「米国で開発された学修評価の基準の作成方

法であり、評価水準である『尺度』と、尺度を満たした場合の『特徴の記述』で構成される。記述により達成水準等が明確化されることにより、他の手段では困難な、パフォーマンス等の定性的な評価に向くとされ、評価者・被評価者の認識の共有、複数の評価者による評価の標準化等のメリットがある」（p. 39）。

- 3) IR とは、データ収集・蓄積、とくに教育機能についての調査・分析、および、大学経営の基礎となる情報・分析の提供をさす（金子，2011）。
- 4) 「第二十五条の二 2 大学は、学修の成果に係る評価及び卒業の認定に当たっては、客観性及び厳格性を確保するため、学生に対してその基準をあらかじめ明示するとともに、当該基準にしたがって適切に行うものとする」という条項である。
- 5) 初等・中等教育分野での評価研究や教育行政では、橋本（1983）の以下の区別にもとづいて「キジュン」に「規準」と「基準」の漢字が当てられてきた。すなわち、「目標規準（*criterion* クライテリオン）——評価・解釈の規準を教育目標においたもの」、「到達基準（*standard* スタンダード）——目標規準を一層具体的に、量的、段階的に示したもので、生徒の業績を直接それに照合することができるところまで具体化して設定されたもの」という区別である。これに従えば、英語の *criterion* には「規準」をあてるべきだということになる。また、田中（2008）は、量的な「基準づくり」では思考力や判断力といった高次の目標を質的に把握するのは難しいために、質的な「基準づくり」の方法論として「ルーブリック（評価指標）」という考え方が登場してきた、という（p. 137）。しかし、ルーブリックは、英語では往々にして“*criteria*”、あるいは“*scoring criteria*”とも言い換えられ、橋本のいう目標規準と到達基準の違いは“*general (generic) criteria*”と“*specific criteria*”といった形で表現されることが多い。このような事情から、本稿では、「規準」と「基準」を区別せず、すべて「基準」で統一することにした。
- 6) 本稿では、標準テストに代わる評価方法として3つの評価方法をあげ、その総称として「オルターナティブ・アセスメント」を用いたが、これとは異なるカテゴライズのしかたもある。たとえば Hart（1994）は、標準テストに代わる評価方法として、代替的評価（*alternative assessment*）、パフォーマンス評価、ホリスティックな評価、真正の評価などをあげた上で、「真正の評価」をこれらの総称として用いている（邦訳 p. 11）。ただし、原著のタイトルが *Authentic Assessment* であるのに対し、邦訳のタイトルは『パフォーマンス評価入門』となっており、ここにもこれらの評価方法の重なりが表れている。なお、ウィギンズ&マクタイ（2012）は、パフォーマンス評価はかなり普及したが真正の評価は今でも稀にしかみられないと現状分析しており（p. ii）、そこでは本稿と同じように、両者が同じカテゴリーレベルに置かれている。
- 7) ここでいう「リフレクションシート」は1コマの授業の終わりに授業をふり返ってその日学んだこと、わかりにくかったことなどを書かせるものであり、ミニッツペーパーと同じく、学生の自己認識を通じて学生の学習成果を間接的に評価し、それによって授業評価の資料とするタイプのものである。
- 8) 尺度が数値ではなく言語で表現される場合もあるが（A, B, C など）、その場合でも順序尺度の性質は備えているので、数値的な尺度とみてさしつかえない。
- 9) 「標準化された（*standardized*）」とはどういう状態を表すのかについて、Wiggins & McGighe（2005）は、「すべての生徒について、運営上の条件と手順が統一されているテストや評価方法を言い表すために使われる用語」とし、「すべてのフォーマルなテストは標準化されている」（邦訳 p. 406）と述べている。本稿でもこの定義に従っている。
- 10) OECD-AHELO については、OECD のウェブサイト参照（http://www.oecd.org/education/highereducationandadultlearning/testingstudentanduniversityperformancegloballyoecdshello.htm#Preliminary_findings）。
- 11) 深堀（2008）によれば、「付加価値」を測定する調査デザインとして、「同一人物の二時点間（入学時と卒業時）の学習成果の差異に着目するパネル調査の方法、新入生と卒業年度生の学習成果の差異に着目するコーホート調査の方法、中等教育段階までの学習成果（初期差異）を統計的に統制する方法、学生が高等教育機関で初めて学習する専門分野での学習成果に注目する方法など」が考えられている、という。
- 12) OECD-AHELO のウェブサイトリンクされている CLA のサンプル課題もこれと同じである（<http://starttest.com/7.0.0.1/programs/clacross/Practice%20Test%20Page.htm>）。

- 13) 前述のように、ウィギンズらはパフォーマンス課題をかなり限定してとらえている。この課題も、彼らの定義にしたがえば、学校や試験という条件下で特定のプロンプトに対し構造化された応答が求められているという点で、「パフォーマンス課題」というよりは、むしろ「アカデミック・プロント」にあたると思われる (Wiggins & McTighe, 2005, 邦訳 p. 185 参照)。
- 14) Benjamin et al. (2009) には採点基準は掲載されておらず、批判的思考・分析的推論・問題解決スキルと文章表現スキルに分けて、評価のポイントと4段階あるいは7段階で評価すべきことが書かれている。本稿で取り上げた採点基準は、CLA のウェブページから引用した (<http://www.collegiatelearningassessment.org/files/CLAScoringCriteria.pdf>)。
- 15) 以下、アルヴァーノ・カレッジの実践については、大学のウェブサイト (<http://www.alverno.edu/>)、とくに Quick Facts、Our Unique Curriculum、The 8 Core Abilities、Alverno's Grading System、Diagnostic Digital Portfolio の各ウェブページ、および、アルヴァーノの教員による著作、Alverno College Faculty (1994)、Loacker & Rogers (2005)、Mentkowski & Associates (2000) を参照した。
- 16) Wiggins & McTighe (2005) や Hart (1994) でも Alverno の取り組みが紹介されている (それぞれ、邦訳 p. 256・332、邦訳 p. 77-78 参照)。
- 17) Project DEEP については、NSSE のウェブページ (http://nsse.iub.edu/_/?cid=70) を参照。
- 18) アルヴァーノのウェブページ (<http://www.alverno.edu/academics/ouruniquecurriculum/the8coreabilities/>) も参照。
なお、最新版では「意思決定における価値判断」は単に「価値判断」となっている。
- 19) 以下の評価組織体制の記述においては、Alverno College Faculty (1994, Chap. 5)、Loacker & Rogers (2005)、ならびに各組織のウェブサイトを参照した。
- 20) VALUE プロジェクトに関する日本での先行研究としては、森 (2011a, 2011b)、吉田 (2011)、笠原 (2011) を参照。
- 21) AAC&U や VALUE プロジェクトについては、AAC&U のウェブサイト参照 (<http://www.aacu.org/about/index.cfm>)。
後述する VALUE ルーブリックも、同じウェブサイトからダウンロードできる。
- 22) このフレーズは、後述する CAAL プロジェクトのウェブページ (<http://aacu.org/caal/>) から取った。
- 23) 補助カリキュラム (cocurriculum) とは、単位は付与されないが、教育目的のために行われる正課外の活動のことである。具体的には、芸術的活動 (演劇、演奏・合唱、ダンスなど)、海外研修、スポーツ、地域活動、インターンシップなどが含まれる。従来は、課外カリキュラム (extra-curriculum) と総称されることが多かったが、学生の学習や成長に対して、正課のカリキュラムと同等の教育的価値をもつとみなされるようになったことで、この名称が使われることが増えている。
- 24) VALUE ルーブリックの訳出と検討は、筆者の担当する大学院授業科目「高等教育方法演習 A」(2012 年前期) において行った。この演習の参加者は、田口真奈 (高等教育研究開発推進センター准教授: 「口頭コミュニケーション」担当)、高橋雄介 (同特定助教: 「探究と分析」担当)、坂本尚志 (同特定助教: 「倫理的推論」担当)、田川千尋 (同特定助教: 「異文化知識・能力」担当)、平山朋子 (教育学研究科博士後期課程 2 年: 「問題解決」担当)、大山牧子 (同左: 「情報リテラシー」担当)、畑野快 (同左: 「生涯学習」担当)、蔣妍 (同左: 「市民参加」担当)、羽山裕子 (同左: 「チームワーク」担当)、山本はるか (同博士後期課程 1 年: 「読解」担当)、齊藤有吾 (教育学研究科修士課程 2 年: 「量的リテラシー」「統合的学習」担当)、蒲雲菲 (高等教育研究開発推進センター研究生: 「創造的思考」担当) であった。なお、「文章コミュニケーション」は筆者が担当し、「批判的思考」はスミス (2012, pp. 10f) を一部改訳し、書式を修正した。なお、各ルーブリックには、ここに訳出した評価基準表以外に、枠組みの説明と用語集がつけられている。
- 25) CAAL プログラムのウェブサイト (<http://aacu.org/caal/>) より抜粋。
- 26) なお、邦訳では「評価の統一性」となっている。西岡 (2005) も参照。

引用文献

- Alverno College Faculty (1994). *Student assessment-as-learning at Alverno College*. Milwaukee, WI: Alverno College Institute.
- 安藤輝次 (2006). 「アルバーノ大学の一般教育カリキュラムの改革」『奈良教育大学紀要』55 巻 1 号 (人文・社会), 65-78 頁.
- Benjamin, R., Chun, M., Hardison, C., Hong, E., Jackson, C., Kugelmass, H., Nemeth, A., & Shavelson, R. (2009). *Returning to learning in an age of assessment: Introducing the rationale of the Collegiate Learning Assessment*. (<http://www.collegiatelearningassessment.org/>)
- 中央教育審議会 (2008). 「学士課程教育の構築に向けて (答申)」参考資料.
- 大学評価・学位授与機構 (2011). 「(高等教育の) 質」「質保証」『高等教育に関する質保証関係用語集 (第3版)』.
(http://www.niad.ac.jp/n_shuppan/package/no9_21_niadue_glossary3_2011_v2.pdf)
- Dewey, J. (Boydston, J. A., Ed.) (1991). *The later works 1925-1953, Vol. 17 (1885-1953, Essays)*. Carbondale, IL: Southern Illinois University Press.
- Douglass, J. A., Thomson, G., & Zhao, C-M. (2012). The learning outcomes race: The value of self-reported gains in large research universities. *Higher Education*, 64, 317-335.
- 深堀聰子 (2008). 「AHELO 国際的な学習成果アセスメント—OECD の試み—」『日本学術新聞』2329 号.
- Gipps, C. V. (1994). *Beyond testing: Toward a theory of educational assessment*. London: Falmer Press. ギップス, C. V. (2001). 『新しい評価を求めて—テスト教育の終焉—』(鈴木秀幸訳) 論創社.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday. ゴッフマン, E. (1974). 『行為と演技—日常生活における自己呈示—』(石黒毅訳) 誠信書房.
- Hart, D. (1994). *Authentic assessment: A handbook for educators*. Menlo Park, CA: Addison-Wesley. ハート, D. (2012). 『パフォーマンス評価入門—「真正の評価」論からの提案—』(田中耕治監訳) ミネルヴァ書房.
- 橋本重治 (1983). 「教育評価基本用語解説」『指導と評価』29 巻 8 号, 38 頁.
- 平山朋子・松下佳代 (2009). 「理学療法教育における自生的 FD 実践の検討—OSCE リフレクション法を契機として—」『京都大学高等教育研究』15 号, 15-26 頁.
- 平山朋子・松下佳代・西村敦 (2012). 「理学療法を主体的に学ぶ—『OSCE リフレクション法』の試み—」小田隆治・杉原真晃 (編著) 『学生主体型授業の冒険 2—予測困難な時代に挑む大学教育—』ナカニシヤ出版, 202-221 頁.
- 金子元久 (2011). 「IR—期待、幻想、可能性—」『IDE 現代の高等教育』528 号, 4-12 頁.
- 笠原千絵 (2011). 「学習成果の評価方法とルーブリックの活用—アメリカの高等教育関連団体と大学におけるインタビュー調査から—」『関西国際大学研究紀要』12 号, 37-46 頁.
- 川嶋太津夫 (2008). 「ラーニング・アウトカムズを重視した大学教育改革の国際的動向と我が国への示唆」『名古屋高等教育研究』8 号, 173-191 頁.
- 経済産業省 (2008). 『今日から始める社会人基礎力の育成と評価』角川学芸出版.
- Loacker, G., & Rogers, G. (2005). *Assessment at Alverno College: Student, program, institutional*. Milwaukee, WI: Alverno College Institute.
- 松下佳代 (2005). 「学生による授業評価—改善と説明責任—」『日本物理学会誌』60 巻 4 号, 297-300 頁.
- 松下佳代 (2007). 『パフォーマンス評価』日本標準.
- 松下佳代 (2010a). 「学びの評価」佐伯胖 (監修)・波部信一 (編) 『「学び」の認知科学事典』大修館書店, 442-458 頁.
- 松下佳代 (2010b). 「〈新しい能力〉概念と教育—その背景と系譜—」松下佳代 (編著) 『〈新しい能力〉は教育を変えるか—学力・リテラシー・コンピテンシー—』ミネルヴァ書房, 1-42 頁.
- 松下佳代 (2012). 「大学カリキュラム」京都大学高等教育研究開発推進センター (編) 『生成する大学教育学』ナカニシヤ出版, 25-57 頁.
- 松下佳代・田口真奈 (2012). 「大学授業」京都大学高等教育研究開発推進センター (編) 『生成する大学教育学』ナカニシヤ出版, 77-109 頁.

松下佳代 (印刷中). 「質的評価の方法」下山晴彦 (編集代表) 『誠信心理学辞典』誠信書房.

Mentkowski, M. & Associates (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco: Jossey-Bass.

溝上慎一 (2009). 「『大学生生活の過ごし方』から見た学生の学びと成長の検討—正課・正課外のバランスのとれた活動が高い成長を示す—」『京都大学高等教育研究』15号, 107-118頁.

文部科学省 (2011). 「大学における教育内容等の改革状況について」(http://www.mext.go.jp/a_menu/koutou/daigaku/04052801/_icsFiles/afieldfile/2011/08/25/1310269_1.pdf).

森利枝 (2011a). 「eポートフォリオによる学修アウトカムの評価」『文部科学教育通信』280号, 16-17頁.

森利枝 (2011b). 「高等教育機関の特性を活かした客観的指標の模索」『文部科学教育通信』282号, 16-17頁.

森利枝 (2012). 「アメリカにおける学習成果重視政策議論のインパクト」深堀聰子 (編) 『学習成果アセスメントのインパクトに関する総合的研究 (国立教育政策研究所プロジェクト研究 研究成果報告書)』106-117頁.

西岡加名恵 (2005). 「教育評価の方法原理」田中耕治 (編) 『よくわかる教育評価』ミネルヴァ書房, 66-75頁.

西岡加名恵・田中耕治 (2009). 「『活用する力』を育てる授業と評価—中学校—パフォーマンス課題とルーブリックの提案—」学事出版.

小野和宏 (2012). 「大学学習法へのパフォーマンス評価の導入—新潟大学歯学部の新年次教育の場合—」大学教育学会第34回大会ラウンドテーブル「学習成果の直接評価に向けて—パフォーマンス評価の可能性—」.

Resnick, L. B. (1989). Introduction. In L.B. Resnick (Ed.) *Knowing, learning and instruction: Essays in honor of R. Glaser*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Resnick, D. P., & Resnick, L. B. (1996). Performance assessment and the multiple functions of educational measurement. In M. B. Kane, & R. Mitchell (Eds.), *Implementing performance assessment: Promises, problems, and challenges*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rhodes, T. (2009a). From the director. *Peer Review*, 11(1), 3.

Rhodes, T. (2009b). The VALUE Project overview. *Peer Review*, 11(1), 4-7.

Rhodes, T. (Ed.) (2010). *Assessing outcomes and improving achievement: Tips and tools for using rubrics*. Washington, DC: Association of American Colleges and Universities.

Schwandt, T. A. (2007). *The Sage dictionary of qualitative inquiry* (3rd ed.). London: Sage Publications. シュワント, T. A. (2009). 『質的研究用語事典』(伊藤勇・徳川直人・内田健監訳) 北大路書房.

Shavelson, R. J. (2010). *Measuring college learning responsibility: Accountability in a new era*. San Francisco: Stanford University Press.

スミス, J. A. (2012). 「ポートフォリオ総論」(森本康彦訳) 小川賀代・小村道昭 (編著) 『大学力を高める—eポートフォリオエビデンスに基づく教育の質保証をめざして—』東京電機大学出版局.

田中耕治 (編) (2005). 『よくわかる教育評価』ミネルヴァ書房.

田中耕治 (2008). 『教育評価』岩波書店.

田中耕治 (編著) (2011). 『パフォーマンス評価—思考力・判断力・表現力を育む授業づくり—』ミネルヴァ書房.

Taylor, C. (1991). *The ethics of authenticity*. Cambridge, MA: Harvard University Press. テイラー, C. (2004). 「『ほんもの』という倫理—近代とその不安—」(田中智彦訳) 産業図書.

東京大学大学経営・政策研究センター (2008). 『全国大学生調査 (第1次報告書)』(<http://ump.p.u-tokyo.ac.jp/crump/cat79/post-11.html#more>)

Vroeijenstijn, A. I. (1995). *Improvement and accountability: Navigating between Scylla and Charybdis (Guide for external quality assessment in higher education)*. Bristol, PA: Taylor & Francis. フローインスティン, A. I. (2002). 『大学評価ハンドブック』(米澤彰純・福留東土訳) 玉川大学出版部.

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development. ウィギンズ, G. & マクタイ, J. (2012). 『理解をもたらすカリキュラム設計—「逆向き設計」の理論と方法—』(西岡加名恵訳) 日本標準.

山田礼子（編）（2009）.『大学教育を科学する—学生教育評価の国際比較—』東信堂.

山田礼子（2012）.『学士課程教育の質保証へむけて—学生調査と初年次教育からみえてきたもの—』東信堂.

吉田文（2009）.「大学生の学習成果の測定をめぐるアメリカの動向」山田礼子（編）（2009）.『大学教育を科学する—学生教育評価の国際比較—』東信堂, 242-263 頁.

吉田武大（2011）.「アメリカにおけるバリューループブリックの動向」『関西国際大学教育総合研究所研究叢書』4号, 1-12 頁.

謝 辞

本稿は、科学研究費基盤研究（C）「深い学習を促すパフォーマンス評価の開発—OSCE-Rを中心に—」（平成24年度～平成26年度、研究代表者：松下佳代、課題番号24530953）の成果の一部である。

資料 1 CLA の採点基準—パフォーマンス課題の場合—

	分析的推論と評価	効果的な文章作成	文章作成の技法	問題解決
6	・ドキュメント・ライブラリで提示されているすべての主要な立論（もしくは分類すべき対象すべてについての重要な特徴）を支持または反駁する事実やアイデアを、ほぼすべて認識している。明白でないことがらについても分析を進めている。 ・ドキュメント・ライブラリから得た大部分の情報について正確に理解している。 ・情報の質について数個の正確な主張を行っている。	・論理的に一貫性のあるやり方で解答を構成することにより、書き手の立論を非常にわかりやすくしている。 ・それぞれの立論に関連する事実やアイデアを妥当な形で包括的に詳述している。情報源が明記されている。	・文法上の決まりを見事に使いこなしている。 ・様々な長さや形式をもつ、よく構成された複雑な文を、一貫して書いている。 ・的確、多様で高度な語彙を巧みに使いこなしている。	・決定と確固とした根拠が示されており、それらは複数の情報源から導き出された信用できるエビデンスに裏打ちされている。他の選択肢も考慮した上で、入手できるエビデンスから最善の決定を提示している。 ※該当する場合： ・結論から論理的に導かれる手順を示している。その意味するところも検討されている。 ・追加調査の必要性を認識し、未解決の問題の大半を扱うための具体的な調査が提案されている。
5	・ドキュメント・ライブラリで提示されているすべての主要な立論（もしくは分類すべき対象すべてについての重要な特徴）を支持または反駁する事実やアイデアを、数個は認識している。 ・ドキュメント・ライブラリの内容の多くを正確に理解している。 ・情報の質について2、3の正確な主張を行っている。	・論理的に一貫性のあるやり方で解答を構成することにより、書き手の立論をわかりやすくしている。 ・それぞれの立論に関連する事実やアイデアを妥当な形で詳述している。情報源が記されている。	・文法上の決まりを非常によく使いこなしている。 ・様々な長さや形式を持つ、よく構成された文を、一貫して書いている。 ・多様で時に高度な語彙を使用しながらアイデアを効果的に伝えている。	・決定と確固とした根拠が示されており、それらは概ね複数の情報源から導き出された信用できるエビデンスに裏打ちされている。他の選択肢は考慮されていない。 ※該当する場合： ・結論から論理的に導かれる手順を示している。その意味するところもほぼ検討されている。 ・追加調査の必要性を認識し、未解決の問題のいくつかを扱うための調査が提案されている。
4	・ドキュメント・ライブラリで提示されているすべての主要な立論（もしくは分類すべき対象すべてについての重要な特徴）を支持または反駁する事実やアイデアを、2、3は認識している。 ・ドキュメント・ライブラリの重要な内容についての正確な理解を短い言葉で示しているが、いくつかの情報は無視している。 ・情報の質についての正確な主張はごくわずかである。	・解答は書き手の立論やそれらの論理を明らかにする方法で整理されているが、明確ではない。 ・事実やアイデアについて妥当な形で詳述を数カ所で行っている。情報源が記されている。	・文法上の決まりをよく使いこなしているが、わずかな誤りがみられる。 ・構成や長さの多様性がいくつかみられる、よく構成された文を書いている。 ・使用されている語彙は明確にアイデアを伝えているが、多様性に欠けている。	・決定とそれを裏づける信用できるエビデンスが示している。信用できるが対立するエビデンスについては、説明できていない場合がある。他の選択肢を無視しようとしている可能性がある。 ※該当する場合： ・結論から論理的に導かれる手順を示している。その意味するところも手短かに検討している。 ・追加調査の必要性を認識している。1つの未解決の問題を扱うための調査が提案されている。
3	・ドキュメント・ライブラリで提示されている数個の立論（もしくは分類すべき対象すべてについての重要な特徴）を支持または反駁する事実やアイデアを、2、3は認識している。 ・重要な情報を無視しているか、もしくは情報を少し誤って解釈している。情報を“そのまま”再述している場合もある。 ・情報の質についての正確な主張はほぼ示されておらず、いくつかの信頼性のないエビデンスを信用できるものとして示している。	・提示されている立論は限定的もしくは不明確である。それぞれの解答について関連性のある情報を示しているが、それらの情報は立論に練りこまれていない。 ・事実やアイデアについて、提示されたわずかな詳述の中のいくつかは妥当である。いくつかの情報源が不明である。	・文法上の決まりを使いこなせているが、小さな誤りが頻繁にみられる。 ・読みやすい文を書いているが、構成や長さの多様性に欠ける傾向がみられる。 ・使用されている語彙はアイデアを伝えるには十分であるが、多様性に欠けている。	・決定を示しているか示唆しており、理由もいくつか示されているが、その根拠は、説明できていないエビデンスによって反駁される可能性がある。 ※該当する場合： ・手順は手短かに示しているが、結論から論理的に導かれない点がいくつか含まれている可能性がある。 ・追加調査の必要性は認識しているかもしれないが、提案された調査はいずれも曖昧であったり、十分には未解決の問題を扱えていない傾向がある。
2	・ドキュメント・ライブラリで提示されている立論（もしくは分類すべき対象すべてについての重要な特徴）を支持または反駁する事実やアイデアを、ほとんど認識していない。 ・ドキュメント・ライブラリの多くを無視、もしくは誤って解釈している。情報を“そのまま”再述している場合もある。 ・情報の質についての正確な主張は全く示されず、信頼性のない情報を信用できるものとして示している。	・提示された立論は限定的、不確か、不明瞭である。提示されている情報は混乱しており、自らの主張を阻害している。 ・事実やアイデアについての詳述はすべて中身がなく、不正確、もしくは信頼できないものである（例えば、書き手の意見のみ述べられているなど）。多くの情報源は不明確である。	・文法上の決まりは使いこなせておらず、多くの小さな誤りやケアレスミスがみられる。 ・常に似たような構成や長さの文を書しており、いくつかの文は理解し難いものである。 ・使用されている語彙は単純なものであり、いくつかは不正確であるか、意味を不明瞭にしている可能性がある。	・決定を示しているか示唆しているが、ほとんど根拠が提示されていないか、信頼性に欠けるエビデンスに頼っている。 ※該当する場合： ・手順は手短かに示しているが、結論から論理的に導かれない点がいくつか含まれている。 ・追加調査の必要性は認識しているかもしれないが、提案された調査はいずれも曖昧であったり、十分には未解決の問題を扱えていない。
1	・ドキュメント・ライブラリで提示されている立論（もしくは分類すべき対象すべてについての重要な特徴）を支持または反駁する事実やアイデアを、全く認識していない。分析のエビデンスを示していない。 ・重要な情報を無視、もしくはひどく誤って解釈している。 ・エビデンスの質についての主張は全く示されておらず、解答は不確かな情報に基づいている。	・説得力のある議論を展開できていない。文章は整っておらず、不明確である。 ・事実やアイデアについての詳述は行われていない。	・文法上の決まりは全く使いこなせておらず、多くの誤りがあるために、解答が理解し難いか、判断するためのエビデンスが不十分にしか示されていない。 ・文は反復的または不完全であり、いくつかは理解し難いものである。 ・使用されている語彙は単純なものであり、いくつかは不正確であるか、意味を不明瞭にしている。	・明解な決定や、決定に至るまでの妥当な根拠が示されていない。 ※該当する場合： ・結論から論理的に導かれる手順を示していない。 ・追加調査の必要性を認識していない、あるいは、未解決の問題を扱う調査を提案していない。

(出典) CLA のウェブページ (<http://www.collegiatelearningassessment.org/files/CLAScoringCriteria.pdf>) より抜粋の上、訳出。

資料2 「能力をベースにしたカリキュラム」の枠組み—8つの能力と6つのレベル—

レベル

上級レベル (専門領域)	6	学問分野・専門職への関わりを反映した、方略、理論、テクノロジーを利用する。	専攻または副専攻の学問分野での枠組みを自主的に適用して、複雑な問題を分析する。	学問分野・専門職の方法や枠組みを適用する（個人的な価値観や見方と統合する、特定の場面に合わせる、問題解決活動を実行する際に自主性や創造性を示す）。	責任ある自己としてイニシアチブを取るために、自分の価値体系をたえず見直し高める。	個人間やグループでの効果的なインタラクションにおいて専門的な目標を達成するためにリーダーシップを発揮する。	特定のグローバル問題に対する理論的・現実的アプローチを創造的・自主的に提案する。	理論を発展させ、生じそうな問題を予想し、対応策を立案するというを試みる。	美的感覚を学問的・職業的・私的生活に組みこむ。
	5	学問分野・専門職の枠組み・理論にあわせてコミュニケーション方略を選択し、修正し、結合する。	枠組みの理解を精緻化し、どんな枠組みが現象を説明するのに適しているかを決定するための規準を特定する。	グループ・プロセスの理解を、協働的問題解決での遂行に転移させる能力を示す。	専門分野や専門職における価値判断の枠組みを用いて、個人・職業人・市民としての重要な問題に取り組む。	複数の学問的枠組みを用いて、行動を解釈したり自分のインタラクションにおける選択をモニターしたりすることで、効果的な専門的インタラクションを一貫して自律的に行う。	理論を用いて、特定のグローバル問題に対する現実的なアプローチを創出する。	市民生活・職業生活において効果的に働き、他者と効果的に協働することで、参加能力を高める。	個人的な美的感覚を育み、表現する。
中級レベル	4	コミュニケーションの諸様式を結びつけ、学問分野の枠組みの中で効果的に統合する。	構造と編成を分析する。	解決策を発展させるために、さまざまなアプローチを自主的に分析し、選択し、利用し、評価する。	特定の学問分野の見方や概念を用いて道徳的判断・決断を特徴づける。	社会的・文化的文脈の認知的理解と自他の行動の感情的要素の認識にもとづいて、グループや個人間での状況において効果的なインタラクションを継続して行う。	多様な学問分野の枠組みを用いて、問題についての自分の判断を明確化する。	計画の有効性の評価規準と行動方略を構築する。	創造的・解釈的プロセスに対する意識を育てる。
	3	さまざまな学問分野で意味を構成するためにコミュニケーション・プロセスを意図的に利用する。	関係性を把握したり、作り出したりする。	学問分野での問題解決プロセスについて、評価や実行を含め、すべての段階・ステップを遂行する。	自分の価値観と社会的文脈との相互関係を分析し、その関係がどこまで使われるかを探る。	社会的・文化的文脈における自他についての注意深い分析にもとづいて、グループや個人間でのインタラクションの有効性を高める。	学問的な概念と枠組みを用いて情報を収集し、グローバルな問題に対してとりうる対応を探究する。	共通の目標を達成するために個人がいかに他者と協同するかという点から、「組織を読む」方法を学ぶ。	学問分野の文脈を統合することによって選択を修正する。
初級レベル	2	コミュニケーションの諸様式に含まれるプロセスやその間の相互作用を認識する。	観察から合理的な推論を導く。	問題にアプローチするために、学問分野での問題解決プロセスの諸要素を実際に使う練習をする。	自分の価値観を行動と結びつけ、そのプロセスの認知的・精神的側面を明確化する。	社会的・文化的文脈においてインタラクションのもたらす感情的・実質的な影響への洞察力を養う。	グローバルな問題を構成している複雑な関係を検討する。	学問的概念を用いて、問題になっていることがらを記述するとともに、情報を収集し、判断を下し、意思決定プロセスに参加するのに必要なスキルを育てる。	芸術的な選択や解釈に対する明確な理由づけを行う。
	1	コミュニケーションの諸様式における自分の長所・短所を認識する。	正確に観察する。	問題にアプローチするステップを明確化することによって、問題解決プロセスを明らかにする。	自他の価値観とそれの引き起こす感情を把握する。	インタラクションにおける自分の行動を自覚し、インタラクションに十全に参加する方法として、分析的な枠組みを認識する。	グローバルな関心事について思考し取り組む能力という点から、自分の知識やスキルを評価する。	自己評価スキルを育て、地域での経験を記述する枠組みを認識し始める。	情報にもとづいた芸術的・解釈的選択をする。
		コミュニケーション	分析	問題解決	意思決定における価値判断	社会的インタラクション	グローバルな視野の発達	効果的な市民参加	美的な関わり

(出典) Loacker & Rogers, 2005, pp. 37-40 より訳出の上、図表化。

資料 3-1 探究と分析 VALUE ルーブリック

定義

探究とは、情報に基づき結論や判断をもたらす論拠（エビデンス）の収集と分析を通じて、論点、関心事、作品（ワーク）について検討する体系立ったプロセスである。分析とは、複雑なテーマや論点についてより良く理解するために、それらをいくつかの部分に分解するプロセスである。

* ペンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ベンチマーク 1
テーマ選択	創造的で焦点化された扱いやすいテーマを選んでいて、重要である可能性があるがこれまでに検討されていない側面を取り上げている。	焦点化して扱いやすいテーマを選んでいて、そのテーマに関連する側面も適切に取り上げている。	扱いやすいテーマを選んでいるが、テーマを絞りすぎていて、そのテーマの関連する側面を見逃している。	一般的で幅が広すぎるために扱いにくいテーマを選んでいる。
これまでの知見・研究・観点	さまざまな観点やアプローチについて述べている、適切なソースからの詳細な情報を総合的に扱っている。	さまざまな観点やアプローチについて述べている、適切なソースからの詳細な情報を示している。	限られた観点やアプローチについてしか述べていない、適切なソースからの情報を示している。	限られた観点やアプローチについてしか述べていない、適切ではないソースからの情報を示している。
デザインのプロセス	方法論や理論的な枠組みのすべての要素が巧みに説明されている。学問領域中からもしくは関連する下位領域から、適切な方法論もしくは理論的な枠組みを総合的に扱うことができている。	方法論もしくは理論的な枠組みにおける重要な要素は適切に説明されているが、より細かい要素は無視されたり説明されていない。	方法論もしくは理論的な枠組みにおける重要な要素が欠けていたり、誤って説明されたり、焦点が絞れていない。	探究のデザインにおいて、方法論もしくは理論的な枠組みに対する誤解をしている。
分析	論拠を整理して総合的に扱い、問題の焦点に関連する洞察に富んだパターン、違い、類似性を明らかにしている。	論拠を整理して、問題の焦点と関連する重要なパターン、違い、類似性を明らかにしている。	論拠は整理されているが、その整理が重要なパターン、違い、類似性を明らかにするには効果的ではない。	論拠は並んでいるが、整理されておらず、問題の焦点と関連していない。
結論	探究で得た結果から、論理的に既知の事柄のことから推定を行った結論を述べている。	ただ単に探究で得た結果に焦点を当てた結論を述べている。その結論は、明確に探究で得た結果から上がってきたものであり、その結果に明確に対応している。	一般的な結論を述べている。結論がとも一般的なので、探究で得た結果の範囲を超えてしまっている。	探究で得た結果から、あいまいで筋が通らず支持されない結論を述べている。
限界点と示唆	適切で裏付けのある限界点と示唆について、洞察に富んだ議論を詳細に行っている。	適切で裏付けのある限界点と示唆について議論を行っている。	適切で裏付けのある限界点と示唆を述べている。	限界点や示唆を述べているが、それらは場合によっては不適切であり裏付けに乏しい。

(注) 訳出は高橋雄介（京都大学高等教育研究開発推進センター特定助教）による。

資料 3-2 批判的思考 VALUE ルーブリック

定義

批判的思考とは、ある意見や結論を受容したり案出したりする前に、問題、アイディア、アーティファクト（人間の手で作り出された物）、出来事などを包括的に探究するという特徴をもつ思考の習慣のことである。

*ペンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ペンチマーク 1
課題説明	批判的に考えるべき課題や問題が、完全に理解するために必要となる重要な情報を伴って、明確に記述され、包括的に説明されている。	批判的に考えるべき課題や問題が明確に記述・説明されており、無用な省略が理解を妨げることがない。	批判的に考えるべき課題や問題についての記述・説明が、未定義の用語を用いていた、あるいは残っていたり、議論の境界が不定であったり、背景が不明だったりする。	批判的に考えるべき課題や問題が、明確に記述・説明されていない。
エビデンス 視点や結論を検討するための情報選択・利用	包括的な分析や統合を行うための情報を、十分な解釈や評価を伴い取得している。専門家の視点を完全に批判的にとらえている。	首尾一貫した分析や統合を行うための情報を、十分な解釈や評価を伴い取得している。専門家の視点を批判的にとらえている。	ある程度の解釈や評価を伴い情報を取得しているが、首尾一貫した分析・統合を行うには不十分である。専門家の視点を、批判することなく、ほとんど事実として取り上げている。	解釈や評価なしに情報を取得している。専門家の視点を、批判することなく事実として取り上げている。
文脈や仮定の影響	自己や他者の仮定を徹底的（体系的・方法的）に分析し、見解を示す際、文脈の重要性を慎重に評価している。	見解を示す際、自己または他者の仮定や重要な文脈を確認している。	仮定のいくつかを批判的にとらえている。見解を示す際、重要ないくつかの関係ある文脈を確認している。自分自身の仮定よりも他者の仮定に対する（あるいはその逆）さらなる配慮が必要である。	現在の仮定に気づきつつあることが見取れる（主張を仮定とみなしているときもある）。見解を示す際、文脈のいくつかを確認し始めている。
学生の見解 (展望、主張または仮説)	想像力に富み課題の複雑さを考慮した具体的な見解（展望、主張または仮説）が示されている。見解（展望、主張または仮説）の限界を認めており、他の視点と統合されている。	課題の複雑さを考慮した具体的な見解（展望、主張または仮説）が示されている。他の視点が、見解（展望、主張または仮説）の内に認められる。	具体的な見解（展望、主張または仮説）が、課題の違う側面を認めている。	具体的な見解（展望、主張、または仮説）が示されているが、それは簡素で自明である。
結論と関連する成果	結論や関連する成果が論理的で、学生の知識に基づいた評価や、優先順位を付けて議論されたエビデンスや展望を提示する能力を反映している。	結論が、反対の視点を含め、さまざまな情報に論理的に結び付けられている。関連する成果が明確に確認できる。	結論が情報に論理的に結び付けられている（望ましい結論に合うように情報が選ばれている）。関連する成果がいくつかが明確に確認できる。	結論が議論された情報のいくつかに矛盾して関連付けられている。関連する成果を簡略化しすぎている。

(出典) スミス (2012, pp. 10f) より抜粋、一部改訳・書式修正。

資料 3-3 創造的思考 VALUE ループリック

定義

創造的思考とは、既存のアイデア、イメージや専門的知識をオリジナルなやり方で結合し統合する能力であり、また、高度な革新、拡散的思考、リスクテキーキングを特徴とする想像的なやり方で思考し、反応し、作業する経験である。

*ベンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ベンチマーク 1
コンピテンスの獲得	省察：その領域の適切な規準を用いて、創造的なプロセスと成果物を評価している。	創造：その領域にとって適切であり、しかもまったく新しい事物・解決方法・アイデアを創造している。	適応：適切な見本を自分自身の様式にうまく適応させている。	模倣（model）：適切な見本をうまく模倣している（reproduce）。
リスクテキーキング	最終成果物となる課題に対して、未だ検証されておらず潜在的なリスクを含むような方向やアプローチを積極的に追求し、最後までやり遂げている。	最終成果物となる課題に対して、新しい方向やアプローチを取り入れている。	課題のガイドラインを超えない範囲で、新しい方向やアプローチを考えている。	課題のガイドラインの内に、厳密に留まっている。
問題解決	論理的で首尾一貫した計画を作るだけでなく、解決方法の重要性を認識し、解決方法の選択理由をはっきり説明することができる。	問題解決のために、複数の選択肢から一つを選択し、論理的で首尾一貫した計画を作っている。	問題解決に対して、あまり受け入れられないアプローチを考慮した上で、却下している。	問題解決のためにただ一つのアプローチを考慮し、利用している。
矛盾の受け入れ	代替的、拡散的、あるいは矛盾をはらんだ見方やアイデアを完全に統合している。	代替的、拡散的、あるいは矛盾をはらんだ見方やアイデアを探索的方法で取り入れている。	代替的、拡散的、あるいは矛盾をはらんだ見方やアイデアをわずかに含めている（価値を認めている）。	代替的、拡散的、あるいは矛盾をはらんだ見方やアイデアを認識している（軽く触れている）。
革新的な思考	新奇な（あるいは独創的な）アイデア・問い・型・成果物を拡張して、新しい知識や境界を超える知識を創造している。	新奇な（あるいは独創的な）アイデア・問い・型・成果物を創造している。	新奇な（あるいは独創的な）アイデア・問い・型・成果物の創造を試みている。	利用できるさまざまなアイデアを再定式化している。
合併・統合・変換	アイデアや解決方法を全く新しい形式に変換している。	アイデアや解決方法を首尾一貫性を持った全体の中に統合している。	アイデアや解決方法を新奇なやり方で関連づけている。	アイデアや解決方法の間に存在するつながりを認識している。

(注) 記出は蒲雲非（京都大学高等教育研究開発推進センター研究生）による。

資料 3-4 文章コミュニケーション VALUE ルーブリック

定義

文章コミュニケーションとは、文章作成（ライティング）によって自分の考えを展開・表現することである。文章コミュニケーションには、多くのジャンルと文体を使って書けるようになることが含まれる。また、多様な文章作成テクノロジーを使って書くことや、テキスト・データ・イメージをうまくミックスさせることなども含まれる。文章コミュニケーション能力はカリキュラム全体を通じて、繰り返し書く経験を重ねることによって発達する。

* ペンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	マイルストーン			ペンチマーク 1
	キャプストーン 4	3	2	
文章作成の文脈と目的 読者・目的や、課題をとりまく状況の考慮を含む	文脈・読者・目的について完璧な理解を示し、それによって、与えられた課題に対応し、作品のあらゆる要素に焦点をあてている。	文脈・読者・目的について適切な理解を示し、与えられた課題（例えば、読者・目的・文脈を結びつけること）に明確に焦点をあてている。	文脈・読者・目的や与えられた課題（例えば、読者の認知や了解事項への気づきを見せ始めること）への自覚を示している。	文脈・読者・目的や与えられた課題（例えば、読者としての授業者や自己の期待）に対し最低限の注意を示している。
内容の展開	適切で関連性があり説得力に富む内容を用いることによって、科目の習得ぶりを示すとともに、書き手の理解したことを伝え、作品全体を形づくっている。	適切で関連性があり説得力に富む内容を用いることによって、学問分野の文脈の中でアイディアを探究し、作品全体を形づくっている。	適切で関連性のある内容を用いることによって、作品の大半を通じて、アイディアを展開・探究している。	適切で関連性のある内容を用いることによって、作品の何カ所かで、シンプルなアイディアを展開している。
ジャンルと学問分野の約束事 特定の形式や学問分野の文章作成に期待される公式・非公式のルール（用語集参照）	特定の学問分野や文章作成課題に関連する広範な約束事（構成、内容、提示、書式、文体選択を含む）に対し、細かい注意を向けうまく遂行している。	特定の学問分野や文章作成課題に関連する重要な約束事（構成、内容、提示、文体選択を含む）を一貫性をもって使用している。	特定の学問分野や文章作成課題にふさわしいものとして、期待されることがら（基本的構成、内容、提示など）に従っている。	基本的構成や提示のしかたについて一貫した体系を使おうとしている。
資料（ソース）と根拠（エビデンス）	当該の学問分野やジャンルにふさわしいアイディアを展開するために、質が高く、信頼でき、関連性のある資料をうまく使いこなしている。	当該の学問分野やジャンルの中に位置づくアイディアを裏づけるために、信頼でき、関連性のある資料を一貫して使っている。	当該の学問分野やジャンルにふさわしいアイディアを裏づけるために、信頼できる（もしくは関連性がある）資料を使おうとしている。	アイディアを裏づけるために、資料を使おうとしている。
構文と技法を繰ること	読み手に明確かつ流暢に意味を伝えることができる格調ある言葉遣いをしていいる。ほとんど全く誤りがない。	読み手に意味を伝える直截的な言葉遣いをしていいる。滅多に誤りがない。	文章に数カ所誤りを含むが、明確に意味を伝える言葉遣いをしていいる。	用語法に誤りがあるために、意味の伝達が妨げられるような言葉遣いをしていいる。

（注）訳出は筆者による。

資料 3-5 口頭コミュニケーション VALUE ルーブリック

定義

口頭コミュニケーションとは、自分の知識を増し、理解を深め、あるいは、聞き手の態度や価値観、信念や、行動を変えることができると、準備された、目的をもったプレゼンテーションのことである。

* ペンチマークレベルのパフォーマンス (1のセル) を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ペンチマーク 1
構造	構造的なパターン (具体的な導入と結論、文節で区切られた本文、つなぎ) が明確で、一貫してみられ、上手く、プレゼンテーションの内容を互いに関連づけられている。	構造的なパターン (具体的な導入と結論、文節で区切られた本文、つなぎ) が明確で、プレゼンテーションの中に一貫してみられる。	構造的なパターン (具体的な導入と結論、文節で区切られた本文、つなぎ) が、プレゼンテーションの中に時々みられる。	構造的なパターン (具体的な導入と結論、文節で区切られた本文、つなぎ) がプレゼンテーションの中にみられない。
言葉	言葉の選び方が、想像力に富んでおり、印象的で、説得力があり、プレゼンテーションの効果を増している。プレゼンテーションにおいて、聴衆に応じた言葉を選んでいる。	言葉の選び方が、思慮深く、プレゼンテーションの効果を、おおよそ支えている。プレゼンテーションにおいて、聴衆に応じた言葉を選んでいる。	言葉の選び方が、平凡で、陳腐であり、プレゼンテーションの効果を、部分的にのみ、支えている。プレゼンテーションにおいて、聴衆に応じた言葉を選んでいる。	言葉の選び方が、不明確で、プレゼンテーションの効果は、最小限にとどまっている。プレゼンテーションにおいて、聴衆に応じた言葉を選んでいる。
話し方	話術 (姿勢、ジェスチャー、アイコンタクト、声の表現) が、プレゼンテーションを説得的にしており、話者が洗練され、自信のあるようにみえる。	話術 (姿勢、ジェスチャー、アイコンタクト、声の表現) が、プレゼンテーションを興味深くしており、話者が落ち着いてみえる。	話術 (姿勢、ジェスチャー、アイコンタクト、声の表現) が、プレゼンテーションを理解可能にしているが、話者が自信なさげである。	話術 (姿勢、ジェスチャー、アイコンタクト、声の表現) が、プレゼンテーションの理解を妨げており、話者が落ち着きなくみえる。
サポート資料	多様なサポート資料 (説明、例、図解、統計、比喩、関係ある権威筋からの引用) が、プレゼンテーションを大いに支持し、あるいは、トピックに対する発表者の信頼性や権威を確立するような情報や分析に対する言及を適切なものとしている。	サポート資料 (説明、例、図解、統計、比喩、関係ある権威筋からの引用) が、プレゼンテーションをおおよそ支持し、あるいは、トピックに対する発表者の信頼性や権威を確立するような情報や分析に対する言及を適切なものとしている。	サポート資料 (説明、例、図解、統計、比喩、関係ある権威筋の引用) が、プレゼンテーションを部分的に支持し、あるいは、トピックに対する発表者の信頼性や権威を確立するような情報や分析に対する言及を適切なものとしている。	サポート資料 (説明、例、図解、統計、比喩、関係ある権威筋の引用) が、プレゼンテーションを最低限支持し、あるいは、トピックに対する発表者の信頼性や権威を確立するような情報や分析に対する言及がなされている。
中心的なメッセージ	中心的なメッセージが説得力をもっている (正確に述べ、適切に繰り返し、記憶に残るようなものであり、強く支持されている)。	中心的なメッセージが明確であり、サポート資料によって一貫性をもっている。	中心的なメッセージは基本的に理解可能であるが、繰り返しがなく、記憶に残るものではない。	中心的なメッセージは推測できるものの、プレゼンテーションにおいて明確には述べられていない。

(注) 訳出は田口真奈 (京都大学高等教育研究開発推進センター准教授) による。

資料 3-6 読解 VALUE ルーブリック

定義

読解とは、「書き言葉との相互作用や関与を通して、意味の抽出と構成を同時に進行過程」である (Snow et al., 2002) (www.rand.org/pubs/research_briefs/RB8024/index1.html より抜粋)。

* ペンチマークレベルのパフォーマンス (1 のセル) を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ペンチマーク 1
理解	文脈や見方、教室で与えられた課題を越える問題、もしくは著者の明確なメッセージを越える問題に対して、テキストが含意している可能性のあるものを認識している (例えば、より広範な問題を認識したり、著者のメッセージや説明に異議を唱えたりする)。	著者のメッセージや態度について、より複雑な推論を行うために、テキストや一般的な背景知識、および (もしくは) 著者の文脈に関する特定の知識を使用している。	テキストの特徴 (例えば、文とパラグラフの構造や論調) が著者のメッセージにいかに関与しているのかを評価している。テキストの文脈や目的について、基礎的な推論を行っている。	テキストが伝達している情報を言い換えて、要約したりすることによって、語彙を適切に捉えている。
ジャンル	特定のテキストに関する一般的なニュアンスにもとづいた期待や読解方略をモニタリングしたり、調整したりしながら、ジャンルの内部もしくはジャンルの超えて、テキストを同定する能力を用いている。	ジャンルやその特徴的な約束事の違いを明確に見分けている。	経験的に、また意図的に、流れに沿ったり逆らったりして読みながら、さまざまなジャンルにわたる読解の経験を省察している。	さまざまな授業の読解課題に対して、省察的ではないにしても生産的な方法で、ジャンルに関する暗黙的な知識を適用している。
テキストとの関係 文脈のなかでテキストの意味を構成すること	さまざまな学問分野の内部で、また学問分野をまたがって、学問的な意味や関連性という点からテキストを評価し、テキストの寄与や重要性にもとづいてその評価を行っている。	重要な問いを提起したり、探究したりするために、また学問的な知識基盤を発展させるために、学問の文脈においてテキストを使用している。	時事的・世界的な知識を構築するという意図や期待をもって、テキストと関わっている。	単位を取得するために、事実と概念を学習したり、正しい答えを見つけようとしたりするという意図と期待をもつて、課題という文脈のなかでテキストにアプローチしている。
分析 部分的に、また全体的にテキストと相互作用を行うこと	テキストや学問分野の内部で、またそれらをまたがって知識や見識を構築するために、観念やテキスト構造といったテキストの諸特徴を関連づけるための方略を評価している。	観念やテキスト構造といったテキストの諸特徴の間の関係を同定し、それらがどのようにテキスト全体の高度な理解を支えているのかを評価している。	効果的な (もしくは効果的でない) 議論や文学的特徴など、テキストの諸部分や諸側面の間の関係を認識し、それがどのようなようにテキスト全体の基本的な理解に寄与しているのかを考慮している。	与えられた課題で提示された問いに応えるために必要なものとして、テキストの諸側面 (例えば、内容、構造、観念間の関係など) を同定する。
解釈 意味の青写真としてのテキストを理解すること	適切な認識論的レンズを用いて読むことができないだけでなく、学問分野やコミュニティの読み手の内部、もしくはそれを超えたところで継続している対話の一部としての読解に参加することができるというエビデンスを示している。	一つの学問分野やあるコミュニケーションの読み手に特有の読解方法や解釈方略について理解したことを、明確に述べている。	読解の目的にもとづいて解釈方略を選択し、目的をもって読んでいくことができる。	課題を明らかにするために、講師のような外部の権威に頼りながら、読解の目的を同定することができる。
読み手の声 テキストに関する学問的な対話に参加すること	学問的な会話を進めたり続けたりするために、自立的な知的、倫理的な性向をもって、テキストについての議論を行う。	進行中の議論を深めたり促したりするために (解釈や質問を通じて) テキストについてくわしく説明する。	テキストについての基本的な理解を共有するのに貢献するようなやり方で、(教室などにおける) 構造化された会話において、テキストについての議論を行う。	著者の意図を保持しつつ、それを課題と関連づけるようなやり方で、テキストについてコメントする。

(注) 訳出は山本はるか (京都大学大学院教育学研究科博士後期課程 1 年) による。

数学的技能 (numeracy)、あるいは量の推論能力 (quantitative reasoning) としても知られている量的リテラシー (quantitative literacy) とは、数値データを扱う際の「思考の習慣 (habit of mind)」、コンピテンシー、抵抗の無さ (comfort) のことである。しっかりと量的リテラシーを持つ学生は、現実味のある文脈や日常生活の状況における広範な量的問題を推論し、解決する能力を持ち合わせている。そういった学生は、量的エビデンスによって裏付けられた洗練された議論をすることができ、さまざまな形式で (言葉、表、グラフ、方程式など適切なものを使って) わかりやすくそうした議論を進めることができる。

*ベンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	キャプブストーン 4	3	マイルストーン	ペンチマーク 1
解釈 数学的形式（例：方程式、グラフ、図、表、言葉）で表された情報を説明する能力	数学的形式で表された情報に関して正確な説明を提供できている。さらにその情報はさらなる深い理解に寄与するような方法から、洞察に富んだ数学的表現に巧みに変換している。	数学的形式で表された情報に関する正確な説明を提供している。（例えば、グラフで示されたトレンドデータを正確に説明できている。）	数学的形式で表された情報に関してある程度正確に説明している。しかしときどき計算や単位などに関連して小さい誤差を示している。（例えば、グラフで示されたトレンドデータが、トレンドの傾向を示す直線の傾きで計算ミスをしている）	数学的形式で表された情報に関して説明を試みている。しかし情報が意味を失っていることに気づいて間違った結論を導いている。（例えば、グラフで示されたトレンドデータが、しかし正の傾向か負の傾向かで混乱するなどとしてトレンドの性質を間違えて解釈している）
表現 適切な情報を様々な数学的形式（例：方程式、グラフ、図、表、言葉）に変換する能力	適切な情報を、さらに進んだ、あるいはさらなる深い理解に寄与するような方法から、洞察に富んだ数学的表現に巧みに変換している。	適切な情報を、適切で望ましい数学的表現に完全に交換している。	情報の変換は完璧であるが、しかしその結果である数学的表現は部分的に適切である。	情報の変換は完璧であるが、しかしその結果である数学的表現は不適切である。
計算	行われた計算は基本的には全て成功しており、問題解決のために十分に理解できている。また、計算プロセスはエレガントである。（明快、簡潔など）	行われた計算は基本的には全て成功しており、問題解決のために十分に理解できている。	行われた計算は正しい解を出すことができている。あるいは問題の完全な解決をするために要求された計算の一部しか示すことができていない。	計算を試みているのだが、しかし正しい解は出せておらず、完全ではない。
応用／分析 分析の限界を認識しながら、データの量的分析をもとに判断して適切な結論を導く能力	データの量的分析を思慮深い判断をもとに使用できている。作品で示された結論は入念で洞察に富んでいる。	データの量的分析を確かな判断のもとで使用できている。作品で示された結論は筋が通っており適切である。	データの量的分析を手際の良い判断（インスピレーションやニュアンスに伴わない）を根拠として使用できたり、作品で示された結論はもっともらしいものである。	データの量的分析を仮の、初歩的な判断のための根拠として使用できている。しかし作品で示された結論には戸惑いを感じられ、疑わしい。
仮説 意見やモデル、データ分析から重要な仮説を構築し、評価する能力	明確に仮説を説明したうえで、さらになぜその仮説が適切なかの説得力のある理論的根拠を提示している。加えて、最終的な結論の信憑性は仮説の正確さによって限定されるという点で自覚が示されている。	明確に仮説を説明したうえで、さらになぜその仮説が適切なかの説得力のある理論的根拠を提示している。	明確に仮説の説明をしている。	仮説の説明を試みている。
コミュニケーション 議論のサボート、あるいは作品の目的における量的エビデンスの目的に説明している（どんなエビデンスが使われているのか、エビデンスの形式がどのようになっているかを指定され、文脈化されているのか）	議論あるいは作品の目的とのつながりの中で量的な情報を使用している。そして効果的なフォオマツットで高いクオリティで詳細に説明している。	議論あるいは作品の目的とのつながりの中で量的な情報を使用している。しかしデータは完全に効果的なフォオマツットで示されていない、あるいは説明のいくつかの部分にムラがある。	量的な情報を使用して目的に効果的に結びつけていない。	どの量的エビデンスが適切であるのかに関して議論が示されているのが、しかし十分な明瞭な数値的サポートが提供されていない。（「たくさん」「少しの」「ますます増える」「小ざら」など、擬似的な量を表す言葉などを使っている）

(注) 訳出は斉藤有吾（京都大学大学院教育学研究科修士課程2年）による。

資料 3-8 情報リテラシー VALUE ルーブリック

定義

情報リテラシーは、情報がいつ必要なのかを知るための能力であり、目の前の問題のために情報を同定し、位置づけし、評価し、効果的で責任のある使用や共有を行うことができる能力である（全米情報リテラシーフォーラム (National Forum on Information Literacy) より借用）。

* ペンチマーク レベルのパフォーマンス (1のセル) を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ペンチマーク 1
必要とされる情報の範囲の決定	リサーチエッセイやテーマの範囲を効果的に限定している。キー概念を効果的に決定できる。選択された情報（ソース）を概念と直接に関連付けるか、直接リサーチエッセイに答えるかしている。	リサーチエッセイやテーマの範囲を十分に限定している。キー概念を決定できている。選択された情報（ソース）を概念と関連付けるか、リサーチエッセイに答えるかしている。	リサーチエッセイやテーマの範囲を不完全にしか限定していない（部分的に欠けており、広すぎたり狭すぎたりしたままである）。キー概念を決定できている。選択された情報（ソース）を部分的に概念と関連付けるか、リサーチエッセイに答えるかしている。	リサーチエッセイやテーマの範囲を限定するのが困難。キー概念を決定するのが困難。選択された情報（ソース）を概念と関連付けておらず、リサーチエッセイにも答えていない。
必要とされる情報へのアクセス	良くデザインされた検索方略で、最も適切な情報ソースを効果的に使用して情報にアクセスしている。	多様な検索方略で、いくつかの関連する情報ソースを使用して情報にアクセスしている。検索を絞り込むための能力を示している。	単純な検索方略を使用して情報にアクセスし、限られたソースや類似したソースから情報を検索している。	手当たり次第に情報にアクセスし、関連性や本質に欠けた情報を検索している。
情報とそのソースの批判的な評価	自己と他者の仮定を徹底的（体系的・方法的）に分析し、見解を示す際には文脈の関連性を慎重に評価している。	見解を示す際に、自己と他者の仮定や、いくつかの関連性のある文脈を同定している。	仮定のいくつかを批判的にとらえている。見解を示す際に、重要ないくつかの関連性のある文脈を同定している。自分自身の仮定よりも他者の仮定に意識を向けている（あるいはその逆）。	現在の仮定に気づきつつあることが見て取れる（主張を仮定とみなしているときもある）。見解を示す際に、文脈のいくつかを同定し始めている。
特定の目的を達成するための、情報の効果的な使用	特定の目的を完全に達成するために、ソースからの情報を、深く理解に、発信し、組織化し、総合している。	ソースからの情報を発信し、組織化し、総合している。意図された目的が達成されている。	ソースからの情報を発信し、組織化することができている。ただし、その情報はまだ、総合されておらず、意図された目的は十分達成されていない。	ソースからの情報を発信することができ。ただし、その情報は断片的であり、不適切に使用されている（文脈に沿わない間違った引用や間違ったパラフレーズなど）。その結果、意図された目的は達成されていない。
情報の倫理的・法的なアクセスと使用	学生は、以下の情報使用方略の全てを正確に使用している（引用や出典の使用、パラフレーズ・要約・引用の選択、もとの文脈にてらして正しいやり方で情報を使用すること、常識と帰属を必要とする「＝誰のアイデアかを記すべき」アイデアとを区別すること）。また、公表された情報、内密な情報、および（もしくは）所有権情報の使用に関する倫理的・法的な制限について十分理解していることを示している。	学生は、以下の情報使用方略の3つを正確に使用している（引用や出典の使用、パラフレーズ・要約・引用の選択、もとの文脈にてらして正しいやり方で情報を使用すること、常識と帰属を必要とする「＝誰のアイデアかを記すべき」アイデアとを区別すること）。また、公表された情報、内密な情報、および（もしくは）所有権情報の使用に関する倫理的・法的な制限について十分理解していることを示している。	学生は、以下の情報使用方略の2つを正確に使用している（引用や出典の使用、パラフレーズ・要約・引用の選択、もとの文脈にてらして正しいやり方で情報を使用すること、常識と帰属を必要とする「＝誰のアイデアかを記すべき」アイデアとを区別すること）。また、公表された情報、内密な情報、および（もしくは）所有権情報の使用に関する倫理的・法的な制限について十分理解していることを示している。	学生は、以下の情報使用方略の1つを正確に使用している（引用や出典の使用、パラフレーズ・要約・引用の選択、もとの文脈にてらして正しいやり方で情報を使用すること、常識と帰属を必要とする「＝誰のアイデアかを記すべき」アイデアとを区別すること）。また、公表された情報、内密な情報、および（もしくは）所有権情報の使用に関する倫理的・法的な制限について十分理解していることを示している。

(注) 訳出は大山牧子（京都大学大学院教育学研究科博士後課程2年）による。

資料 3-9 チームワーク VALUE ルーブリック

定義

チームワークとは、個々のチームのメンバーの統制のもとに行われる行動である（チームの課題に対する努力、他者との関わり方、チームで議論を行う際の貢献の質と量）。

* ベンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ベンチマーク 1
チームの話し合いへの貢献	代替的な考えや提案の長所を明確にすることで、チームが前に進むのを助ける。	他者の考えに基づいた代替的な解決法や行動計画を提案する。	グループの作業を前進させるために新たな示唆を与える。	考えを共有するが、グループの作業を前進させはしない。
他のメンバーがチームに貢献することの促進	誰かが参加していない時には、そのことに気付き参加を促すとともに、他のメンバーの貢献を建設的に積み重ねたり統合したりすることによって、チームのメンバーが話し合いに貢献することを促進する、というやり方で、彼らをチームの活動に参加させる。	他のメンバーの貢献を建設的に積み重ねたり統合したりすることによって、チームのメンバーが話し合いに貢献することを促進する、というやり方で、彼らをチームの活動に参加させる。	他のメンバーの見方について言いなおし、さらに（もしくは）それを明確にするような質問を行うことにより、チームのメンバーが話し合いに貢献することを促進する、というやり方で、彼らをチームの活動に参加させる。	発言の順番を回したり、他者の意見を遮ることなく傾聴することで、チームのメンバーをチームの活動に参加させる。
チームの話し合い以外の場での個人的貢献	決められた課題を全て締め切りまでに完成させる。そして、その成果は徹底的かつ包括的でありプロジェクトを前進させるものである。また、チームのメンバーが与えられた課題を自分と同レベルの卓越性で完成させられるように率先して助ける。	決められた課題を全て締め切りまでに完成させる。そして、その成果は徹底的かつ包括的であり、プロジェクトを前進させるものである。	決められた課題を全て締め切りまでに完成させる。そして、その成果はプロジェクトを前進させるものである。	決められた課題を全て締め切りまでに完成させる。
建設的なチームの雰囲気醸成	以下全てを行うことで、建設的なチームの雰囲気を支える。 ・礼儀正しく建設的なコミュニケーションを行うことで、チームのメンバーを丁重に扱う。 ・チームやその作業に対する肯定的な態度を伝えるため、肯定的な話し方や書き方の調子、表情、および（もしくは）ボディランゲージを用いる。 ・課題の重要さや、それを遂行するチームの能力に対する自信を表明すること、チームメンバーを動機づける。 ・チームのメンバーに対して、支援および（もしくは）励ましを与える。	以下全ての3つを行うことで、建設的なチームの雰囲気を支える。 ・礼儀正しく建設的なコミュニケーションを行うことで、チームのメンバーを丁重に扱う。 ・チームやその作業に対する肯定的な態度を伝えるため、肯定的な話し方や書き方の調子、表情、および（もしくは）ボディランゲージを用いる。 ・課題の重要さや、それを遂行するチームの能力に対する自信を表明すること、チームメンバーを動機づける。 ・チームのメンバーに対して、支援および（もしくは）励ましを与える。	以下全ての2つを行うことで、建設的なチームの雰囲気を支える。 ・礼儀正しく建設的なコミュニケーションを行うことで、チームのメンバーを丁重に扱う。 ・チームやその作業に対する肯定的な態度を伝えるため、肯定的な話し方や書き方の調子、表情、および（もしくは）ボディランゲージを用いる。 ・課題の重要さや、それを遂行するチームの能力に対する自信を表明すること、チームメンバーを動機づける。 ・チームのメンバーに対して、支援および（もしくは）励ましを与える。	以下のうちの1つを行うことで、建設的なチームの雰囲気を支える。 ・礼儀正しく建設的なコミュニケーションを行うことで、チームのメンバーを丁重に扱う。 ・チームやその作業に対する肯定的な態度を伝えるため、肯定的な話し方や書き方の調子、表情、および（もしくは）ボディランゲージを用いる。 ・課題の重要さや、それを遂行するチームの能力に対する自信を表明すること、チームメンバーを動機づける。 ・チームのメンバーに対して、支援および（もしくは）励ましを与える。
【チーム内での】対立への対応	破壊的な対立に直接言及し、チーム全体の結びつきや将来的な効果を高めるやりかたで、その対立に取り組み解決することを経営的に助ける。	対立を認識し、把握した上で、それに取り組む続ける。	対立から離れて、共通の基盤や当座の課題に目を向け直す。	代替的な視点、考え、意見を受動的に受け入れる。

（注）訳出は羽山裕子（京都大学大学院教育学研究科博士後期課程2年）による。

資料 3-10 問題解決 VALUE ルーブリック

定義

問題解決とは、オープンエンドな問題に答えるための方略、あるいは望ましい目標を達成するための方略を、デザインし、評価し、そして実行する過程である。

*ベンチマークレベルのパフォーマンス (1 のセル) を満たさない作品事例にはゼロを割り当てること。

	キャップストーン 4	3	マイルストーン 2	ベンチマーク 1
問題の定義	関連性のあるあらゆる文脈要因の根拠をもつて、明確で洞察に富んだ問題記述を構成する能力を示している。	関連性のある大半の文脈要因の根拠をもつて問題記述を構成する能力を示しており、問題記述も適切である。	関連性のある大半の文脈要因の根拠をもつて問題記述を構成する能力を示し始めるが、問題記述は表面的である。	問題記述や関連する文脈要因を同定する際、限定された能力を示している。
方略の同定 (identify)	特定の文脈においてあてはまる、多様な問題解決アプローチを同定している。	特定の文脈においてあてはまる、いくつかの問題解決アプローチを同定している。	特定の文脈においてあてはまる一つの問題解決アプローチを同定している。	特定の文脈においてあてはまらない一つ以上の問題解決アプローチを同定している。
解決／仮説の提案	問題についての深い理解力を示す一つ以上の解決／仮説を提案している。解決／仮説は、文脈要因だけでなく、問題の倫理的・論理的・文化的次元のすべてについても配慮している (sensitive)。	問題についての理解力を示す一つ以上の解決／仮説を提案している。解決／仮説は、文脈要因だけでなく問題の倫理的・論理的・文化的次元の一つについても配慮している。	問題の特定の文脈要因を扱うために個別にデザインされた解決／仮説よりむしろ「画一的な」一つの解決／仮説を提案している。	漠然としていたり、間接的に問題記述を扱っているにすぎないがために評価することが難しい解決／仮説を提案している。
採りうる解決の評価	解決の評価は深く洗練されており (例えば、徹底的にいて洞察に満ちた説明を含む)、以下の点 (問題の変遷を考える、論理／推論を検討する、解決の実行可能性を検討する、解決の影響力を比較考量する) を深く完全に含んでいる。	解決の評価は適切であり (例えば、徹底的した説明を含む)、以下の点 (問題の変遷を考える、論理／推論を検討する、解決の実行可能性を検討する、解決の影響力を比較考量する) を含んでいる。	解決の評価は簡素であり (例えば、深さに欠けた説明)、以下の点 (問題の変遷を考える、論理／推論を検討する、解決の実行可能性を検討する、解決の影響力を比較考量する) を含んでいる。	解決の評価は表面的で (例えば、おおよっぱで浅い説明)、以下の点 (問題の変遷を考える、論理／推論を検討する、解決の実行可能性を検討する、解決の影響力を比較考量する) を含んでいる。
解決の実行	問題の多種多様な文脈要因を、徹底的に深く扱うというやり方で、解決を実行している。	問題の多種多様な文脈要因を、表面的に扱うというやり方で、解決を実行している。	問題記述を扱ってはいるが、関連する文脈要因を無視するやり方で、解決を実行している。	問題記述を直接扱わないやり方で解決を実行している。
結果の評価	次の作業に必要なことについて、十分、具体的に考えて、定義された問題と関わる結果を吟味している。	次の作業に必要なことについて、多少考え、定義された問題と関わる結果を吟味している。	次の作業に必要なことについて、あるとしてもごくわずかだけ考え、定義された問題の観点から結果を再考している。	次の作業に必要なことについて、まったく考えずに、定義された問題の観点から、表面的に結果を再考している。

(注) 訳出は平山朋子 (京都大学大学院教育学研究科博士後期課程 2 年) による。

資料 3-11 市民参加 VALUE ルーブリック

定義

市民参加とは、「我々のコミュニティの市民生活に変化をもたらすために活動し、その変化をもたらすための知識・スキル・価値観・動機づけを組みあわせながら伸ばしていくこと」である。それは、政治的および非政治的なプロセスを通して、あるコミュニティにおける生活の質（クオリティ・オブ・ライフ）を改善すること」である（*Civic Responsibility and Higher Education* Thomas Ehrlich ed., Oryx Press, 2000, Preface, p.vi より抜粋）。加えて、市民参加には、個人的には生活を豊かにし、社会的にはコミュニティにとって役立つという意味において個人的・公的な重要性をもつ活動に参加するということ行為も含まれる。

* ベンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	キャップストーン 4			マイルストーン 3		2	ベンチマーク 1
コミュニティと文化の多様性	コミュニティや文化の多様性の中で活動し、そこから学ぶことによって、自分の態度や信念を調整しているというエビデンスを示している。他者に多様性への参加を促している。			自分の態度や信念が他の文化やコミュニティの態度や信念とどのような点で異なっているかについて熟考している。コミュニティや文化の多様性から学習できることがらに対して好奇心を示している。		自分の態度や信念は他の文化やコミュニティの態度や信念とは異なっているという自覚をもっている。コミュニティや文化の多様性から学習できるところとがらに対して好奇心をわすかながら示している。	片寄った観点から個人としての態度や信念を表している。コミュニティや文化の多様性から学習できるところとがらに対し、無関心であるか抵抗している。
知識の分析	自分自身の学業／フィールド／学問分野から得た知識（事実、理論など）を、市民参加や市民生活・政治・行政機関への参加とつなぎ、さらにその知識を拡張している。			自分自身の学業／フィールド／学問分野から得た知識（事実、理論など）を、市民参加や市民生活・政治・行政機関への参加とつなぎ始めている。		自分自身の学業／フィールド／学問分野から得た知識（事実、理論など）のうち、市民参加や市民生活・政治・行政機関への参加と関連性のあるものを識別し始めている。	自分の学業／フィールド／学問分野から得た知識（事実、理論など）のうち、市民参加や市民生活・政治・行政機関への参加と関連性のあるものを識別し始めている。
市民のアイデンティティとコミットメント	市民参加活動での経験のエビデンスを提示し、市民のアイデンティティについての強く明確な感覚や公的活動への持続的なコミットメントに関して、自分が何を学んだかを述べている。			市民参加活動での経験のエビデンスを提示し、市民のアイデンティティやコミットメントの感覚が高まるにつれて、自分が自分自身について何を学んだかを述べている。		市民参加活動への参加は、市民のアイデンティティの感覚からではなく、期待やコースの要件から生じているということが、エビデンスによって示されている。	市民参加活動における自分の経験のエビデンスをほとんど提示しておらず、経験と市民のアイデンティティを結びつけていない。
市民のコミュニケーション	さらなる市民活動とのつながりを作るために、効果的に表現し、聞く、他者に合わせるためのコミュニケーション方略を作りあげている。			市民的文脈で効果的にコミュニケーションを行い、〈表現する、聞く、他者の見方にもとづいてアイデアやメッセージを作りかえる〉のすべてを行う能力を示している。		市民的文脈でコミュニケーションを行い、〈表現する、聞く、他者の見方にもとづいてアイデアやメッセージを作りかえる〉のすべてを行う能力を示している。	市民的文脈でコミュニケーションを行い、〈表現する、聞く、他者の見方にもとづいてアイデアやメッセージを作りかえる〉のうちの一つを行う能力を示している。
市民の行動と省察	複雑な（あるいは多様な）市民参加活動について、自立的な経験や、チームリーダーシップのイニシアチブがあることを示す。その際に、自分の行動の目標や遂行についての省察的な視点や分析を伴っている。			市民の行動の自立的な経験やチームリーダーシップがあることを示す。その際に、自分の行動の目標や遂行についての省察的な視点や分析を伴っている。		市民性に焦点化した行動（civically focused actions）にはつきりとした形で参加したことがある。また、これらの行動が、どのように個人やコミュニティのためになるのかを省察（あるいは論述）し始めている。	いくつかの市民的活動を体験したことはあるが、その目標や影響に関する理解の内面化や将来の活動へのコミットメントが乏しい。
市民的な文脈／組織（structure）	市民的な目標を達成するために、コミュニティの文脈や組織にまたがって（あるいはその内部で）協調的に働く能力やコミットメントがあることを示す。			市民的な目標を達成するために、コミュニティの文脈や組織の内部で積極的に働く能力やコミットメントがあることを示す。		市民的な文脈や組織に参加するための計画的やり方を識別する経験があることを示す。	市民的な文脈や組織で体験を行い、何が適しているのかを試している。

（注）訳出は、蔣妍（京都大学大学院教育学研究科博士後期課程2年）による。

資料 3-12 異文化知識・能力 VALUE ルーブリック

定義

異文化知識・能力とは、「多様な文化的文脈において効果的で適切なインタラクションを行うのを下支える認知的・情動的・行動的なスキルと特質の集合」(Bennett, J. M. 2008. Transformative training: Designing programs for culture learning. In *Contemporary leadership and intercultural competence: Understanding and utilizing cultural diversity to build successful organizations*, ed. M. A. Moodian, 95-110. Thousand Oaks, CA : Sage) のことを言う。

* ペンチマークレベルのパフォーマンス (1 のセル) を満たさない作品事例にはゼロを割り当てること。

	キャップストーン 4	3	マイルストーン 2	ペンチマーク 1
知識 文化的自覚	自身の文化的規則や偏見についての洞察をはっきりと述べている。(例えば複雑性を追求する。すなわち、自分の経験がどのようによつてこれらの規則を形づくってきたのか、また、文化的偏見をどう認識し、それにどう応えるのか、ということに自覚的であり、結果的に自己描写に変化をもたらす。)	自身の文化的規則や偏見について新たな見方を認識している。(例えば同一であることを期待しない、すなわち新たな視点を提供する複雑性を受け入れるなど。)	自身の文化的規則や偏見を識別している。(例えば、自身の文化的グループで共有されている規則を強く好み、他者の中に自分と同一であることを求めるなど。)	自身の文化的規則や偏見について (自身の文化的グループで共有されているものでも) 最小限の自覚を示している。(例えば、他者との文化的差異を識別し不快に感じるなど。)
知識 文化的な世界観の枠組みについての知識	歴史、価値観、政治、コミュニケーションスタイル、経済、あるいは信条や実践に關し、自分と異なる文化の成員にとって重要な要素が複雑であることを、洗練された理解を示している。	歴史、価値観、政治、コミュニケーションスタイル、経済、あるいは信条や実践に關し、自分と異なる文化の成員にとって重要な要素が複雑であることを、十分に理解を示している。	歴史、価値観、政治、コミュニケーションスタイル、経済、あるいは信条や実践に關し、自分と異なる文化の成員にとって重要な要素が複雑であることを、部分的な理解を示している。	歴史、価値観、政治、コミュニケーションスタイル、経済、あるいは信条や実践に關し、自分と異なる文化の成員にとって重要な要素が複雑であることを、表面的な理解を示している。
スキル 共感	自身の世界観、および一つ以上の世界観の見方から異文化経験を解釈するとともに、異なる文化を持つグループの感情を認識した支持的なやり方で行動できる能力を示している。	一つ以上の世界観の知的・感情的側面を認識し、インタラクションにおいて一つ以上の世界観を用いることがある。	他の文化的見方の構成要素を識別しているが、あらゆる状況において自身の世界観で反応している。	他者の経験を見てはいるが、自身の文化的世界観を通してそれを行っている。
スキル 言語的・非言語的コミュニケーション	言語的・非言語的コミュニケーションにおける文化的差異についての複雑な理解を明確に述べている (例えば、さまざまな文化において人々がコミュニケーションする際に、どの程度身体的コンタクトを用いるか、あるいは、どの程度直接的／間接的な意味や明白／暗黙の意味を用いるかについて理解しているということを示す)。また、それらの差異に基づきながら共通理解を得るべくうまく交渉することができる。	言語的・非言語的コミュニケーションにおける文化的差異を認識し、それに参加している。また、それらの差異に基づきながら共通理解を得るべく交渉しはじめている。	言語的・非言語的コミュニケーションにおける文化的差異をいくつかに識別することができる。誤解と起こるものにはこれらに基いて、まだ共通理解を得るべく交渉しているが、まだ共通理解を得るべく交渉することはできない。	言語的・非言語的コミュニケーションにおける文化的差異について最低限レベルの理解を持っている (にすぎない)。共通理解を得るべく交渉することができていない。
態度 好奇心	他の文化についての複雑な問いを投げかけ、これらの問いに対し多文化的な見方を投影した答えを探し出し、それをはっきりと述べている。	他の文化について深い問いを投げかけ、これらの問いに対する答えを探し出している。	他の文化について単純で表面的な問いを投げかけている。	他の文化についてもっと学びたいという最低限の興味を述べている。
態度 オープンさ	文化的に異なる他者とのインタラクションを自分から始めて、それを発展させている。文化的に異なる他者とのインタラクションについて価値判断を保留している。	文化的に異なる他者とのインタラクションを自分から始めている。文化的に異なる他者とのインタラクションについて価値判断を保留し始めている。	文化的に異なる他者とのインタラクションの (すべてではないにせよ) 大半において、オープンさを表している。文化的に異なる他者とのインタラクションについて価値判断を保留するのは困難であるが、自身の判断に気付けており変えようとする意志を表している。	文化的に異なる他者とのインタラクションを受け入れることができる。文化的に異なる他者とのインタラクションについて価値判断を保留するのは困難であり、自身の判断に気付けていない。

(注) 訳出は田川千尋 (京都大学高等教育研究開発推進センター特定助教) による。

資料 3-13 倫理的推論 VALUE ルーブリック

定義

倫理的推論とは人間の行動の正誤に関する推論である。学生に要求されることは、学生自身の倫理的価値観と問題の社会的文脈を評定できること、さまざまな状況における倫理的な問題点を認識できること、どれほど異なるさまざまな倫理的視点が、倫理的ディレンマに対して適用されうるかについて考えられること、そしてさまざまな行動のもたらす選択肢の結果について考察できること、である。学生の倫理的自己認識の発展は、倫理的決断のスキルを行使し、さまざまな立場や倫理的な問題点をいかに記述し、分析するかを学ぶことで行われる。

* ペンチマークレベルのパフォーマンス (1 のセル) を満たさない作品事例にはゼロを割り当てること。

	マイルストーン			ベンチマーク 1
	キャプストーン 4	3	2	
倫理的自己意識	学生は中核的信念とその由来について細部にわたって議論し、分析している。そして議論はより深く、明確になっている。	学生は中核的信念とその由来について細部にわたって議論し、分析している。	学生は中核的信念とその由来の両方について述べている。	学生は自分の中核的信念を言明するか、あるいはそうした信念の由来について明確に述べているが、双方を行うことはない。
さまざまな倫理的視点／概念の理解	学生は自分の用いている主要な理論について名前を挙げられ、その要点を示すことができ、理論あるいは諸理論の細部について明確に説明している。	学生は自分の用いている主要な理論について名前を挙げられ、その要点を示すことができ、理論あるいは諸理論の細部について説明を試みているが、幾分の不正確さを伴っている。	学生は自分の用いている主要な理論について名前を挙げられる。また、その要点を示すことはできる。	学生は自分の用いている主要な理論について、ただその名前を挙げるのみである。
倫理的問題点の認識	倫理的な諸問題点が複雑で重層的な（曖昧な）文脈において提示された時に、学生はそれらの問題点を認識でき、それとともに、それらの間の交差関係を把握できる。	倫理的な諸問題点が複雑で重層的な（曖昧な）文脈において提示された時に、学生はそれらの諸問題点を認識でき、あるいは、それらの間の交差関係を把握できる。	学生は基本的かつ明白な倫理的問題を認識でき、（不完全ながら）それらの間の複雑さや相互関係を把握できる。	学生は基本的かつ明白な倫理的な問題点を認識できるものの、複雑さや相互関係を把握できない。
倫理的視点／概念の適用	学生は倫理的視点／概念を、倫理的問いに対して（新たな例を取って）自在にかつ正確に適用でき、適用がもたらす結果について十分に考慮できる。	学生は倫理的視点／概念を、倫理的問いに対して（新たな例を取って）自在にかつ正確に適用できるものの、この適用に固有の様々な結果については考慮しない。	学生は倫理的視点／概念を倫理的問いに対して、自在に（新しい例に）適用できるが、その適用は不正確である。	学生は倫理的視点や概念をある倫理的問いに対して、論拠（教室で、ある集団内での事例の使用、あるいは条件を付けた場面設定）を用いつつ適用できる。しかし、倫理的視点／概念を自在には（新しい例に）適用できない。
さまざまな倫理的視点／概念の評価	学生は自分の立場を明らかにし、それに対する反論、その前提や含意について述べることができ、反論や異なる倫理的視点／概念の前提や含意に対して自分の主張を合理的に擁護できる。学生の擁護は十分かつ効果的である。	学生は自分の立場を明らかにし、それに対する反論やその前提、含意について述べることができるが、それらに対する学生の解答は不十分である。	学生は自分の立場を明らかにし、それに対する反論や異なる倫理的視点／概念の前提や含意について述べることはできるが、それらに答えることはできない（そして、結局のところ、反論、前提そして含意は学生によって細分化されてしまい、学生の立場に影響を与えない）。	学生は自分の立場を明らかにするものの、異なる視点／概念への反論、それらの前提や限定について述べることはできない。

(注) 訳出は坂本尚志（京都大学高等教育研究開発推進センター特定助教）による。

資料 3-14 生涯学習の基礎とスキル VALUE ルーブリック

定義

生涯学習とは「知識、スキル、そしてコンピテンシスの改善を行うという目的を継続的に引き受ける全ての目的学習活動」である。高等教育における試みとは、大学に在籍している間に、このルーブリックに記述されている特定の傾向やスキルを発達させることで、このタイプの学習者になるよう学生を準備させることである (The European Commission, 2000, Commission staff working paper: A memorandum on lifelong learning より抜粋)。

* ペンチマークレベルのパフォーマンス (1のセル) を満たさない作品事例にはゼロを割り当てること。

	マイルストーン			ペンチマーク 1
	キャップストーン 4	3	2	
好奇心	トピックを深く探究し、十分な自覚やほとんど知られていない情報を提供し、テーマに強い興味を示す。	トピックを深く探究し、洞察や情報を生み出し、テーマに興味を示す。	トピックの探究に深さのエビデンスが見られ、洞察や情報をとどこどころ提供し、テーマに対して軽い興味を示す。	トピックを浅いレベルで探究し、きわめて基礎的な事実をこえた洞察や情報をほとんど提供せず、テーマに対して低い興味を示す。
自発性	求められたワークを仕上げ、知識、スキル、能力を拡大する機会を生み出し、追求する。	求められたワークを仕上げ、知識、スキル、能力を拡大する機会を同定し、追求する。	求められたワークを仕上げ、知識、スキル、能力を拡大する機会を同定する。	求められたワークを仕上げる。
自立性	教育的な興味と追求は、授業で要求されること以外にも存在し、豊かに展開している。知識や経験が、自立的に追求される。	授業で要求されること以上に、しっかりと付加的な知識を追求し、自立的な教育経験を能動的に追求する。	授業で要求されること以上に、付加的な知識を追求し、自立的な教育経験を追求することに興味を示す。	授業で要求されることの先を考え、知識を自立的に追求することに興味を示す。
転移	以前の学習を明確に参照し、知識やスキルを、新たな状況での理解やパフォーマンスの提示に革新的 (新しい、創造的) に応用する。	以前の学習を参照し、知識やスキルを、新たな状況での理解やパフォーマンスの提示に応用し、エビデンスを提示する。	以前の学習を参照し、知識やスキルを、新たな状況での理解やパフォーマンスの提示に応用しようとする。	以前の学習を曖昧に参照するものの、知識やスキルを、新たな状況での理解やパフォーマンスの提示に応用することはない。
省察	以前の学習 (授業内・外での過去の経験) を深く再検討し、教育的・日常的出来事についての見方の重要な変化を示している。それらは、長期間にわたって知識、成長、成熟を拡張するための基盤を提供する。	意味を十分明確化したり、教育的・日常的出来事について以前より広い見方を示したりしながら、以前の学習 (授業内・外での過去の経験) を深く再検討している。	意味を多少明確化したり、教育的・日常的出来事について以前よりいりから広い見方を示したりしながら、以前の学習 (授業内・外での過去の経験) をある深さで再検討している。	意味を明確化したり、教育的・日常的出来事について以前より広い見方を示したりすることがないままに、以前の学習 (授業内・外での過去の経験) を浅いレベルで再検討している。

(注) 訳出は畑野快 (京都大学大学院教育学研究科博士後期課程2年) による。

資料 3-15 統合的学習 VALUE ルーブリック

定義

統合的学習とは、学生がアイデアと経験との間の単純なつながりを作ることから、キャンパス内やキャンパスを超えた新しく複雑な状況に対して学習を統合し転移させることにいたるまで、カリキュラムと補助カリキュラムをまたいで形成する理解と性向である。

* ペンチマークレベルのパフォーマンクス（1のセル）を満たさない作品事例にはゼロを割り当てること。

	キャプストーン 4	3	マイルストーン 2	ベンチマーク 1
経験とのつながり 関連する経験とアカデミックな知識とをつなげる	学問分野での理解を深め、そして自分の見方を広めるために、授業外での諸経験（生活経験や、インターンシップや海外旅行などのアカデミックな経験も含む）の間のつながりを意義深く統合している。	学問分野の概念／理論／枠組みを明確にするために、様々な文脈（家族生活・芸術への関与・市民参加・仕事経験）から生活経験の例を効果的に選び、発展させている。	生活経験とアカデミックな知識を、差異と類似性を判断するために比較し、自分とは異なる見方も認めている。	生活経験と、自分の興味との類似性や関連性が認められるアカデミックな文章やアイデアとの間のつながりを同定している。
学問分野とのつながり 学問分野や見方をまたぐつながりがわかる（をつくる）	自発的に、多数の部分から全体像を作り出している（統合している）。あるいは、2つ以上の学問分野や見方からの事例／事実／理論を結合させて結論を導き出している。	自発的に、2つ以上の学問分野や見方からの事例／事実／理論の間につながりを作っている。	（そうするように）促されたときに、2つ以上の学問分野や見方からの事例／事実／理論の間につながりをつくっている。	（そうするように）促されたときに、2つ以上の学問分野や見方からの事例／事実／理論を提示している。
転移 ある状況で得たスキルや能力、理論または方法論などを新しい状況に適用し、応用する	難しい問題を解決するため、あるいは複雑な課題を探索するために、ある状況で得た基本的な方法／スキル／能力／理論／方法論を、新しい状況においてオリジナルな方法で適用し応用している。	問題（problem）を解決するため、あるいは問題点（issues）を探索するために、ある状況で得た基本的な方法／スキル／能力／理論／方法論を、新しい状況で適用し応用している。	問題と問題点（problems and issues）の理解に貢献するように、ある状況で得た基本的な方法／スキル／能力／理論／方法論を、新しい状況で用いている。	ある状況で得た基本的な方法／スキル／能力／理論／方法論を新しい状況で用いている。
統合的コミュニケーション	言語と意味、思考、表現の相互依存関係を明らかにしながら意味を増すようなやり方で、型／言語／グラフ（あるいは他の視覚的表現）を選択することによって、アサインメントをやり遂げている。	何を伝えるか（内容）といかに伝えるか（形式）を、目的や聴衆を意識しながら明確につなげるために、型／言語／グラフ（あるいは他の視覚的表現）を選択することによって、アサインメントをやり遂げている。	何を伝えるか（内容）といかに伝えるか（形式）を基本的なやり方でつなげるような、型／言語／グラフ（あるいは他の視覚的表現）を選択することによって、アサインメントをやり遂げている。	アサインメント（例：小論文、ポスター、ビデオ、パワーポイントのプレゼンテーションを作るなど）を適切な形式でやり遂げている。
省察と自己評価 これまでの経験にもとづいて新しく挑戦的な文脈に対応しながら、学習者としての自己意識の成長を示している（自己評価や省察的・創造的作業のなかで明らかにされるであろう）	多くの多様な文脈をまたいで存在しているであろう将来の自己を思い描いている。（さらに、おそらくは過去の経験にもとづいて計画を立てている。）	長期間にわたる自分の学習の変化を評価しており、複雑な文脈要因（両義性やリスクに取り組む、フラストレーションに対処する、倫理的枠組みを考慮するなど）を認識している。	（特定のパフォーマンスや出来事の中での）強みと課題を明確に表現し、（自己認識を次第に高めることで）別の文脈での有効性を高めている。	自分のパフォーマンスを、成功か失敗かという一般的な記述語で記述している。

（注）訳出は斉藤有吾（京都大学大学院教育学研究科修士課程2年）による。