

発散的課題の学習評価における教師の力量形成

——ループリックは助けか足枷か——

石田 智敬

はじめに

近年、教育における学力観と評価観が大きく転換している。学力観に関していえば、事実的知識や個別的技能の習得よりも、知識や技能を使いこなすような力（認知的に複雑なコンピテンス）をどのように育むかに教育の焦点が向けられている。このような学力は、「高次の学力」と呼ばれ、新しい知の世界を切り拓いたり、複雑な問題解決を遂行したり、様々な情報やデータを分析・評価したりする上で重要となる¹。

こうした学力観の転換は、これまでの評価の考え方を改めることを要求している。従来、特に中等教育段階の成績評価は、中間・期末考査として実施される客観テストの得点に大きな比重を置くことが通例である。ただし、客観テストは、比較的単純な知識や技能の習得の評価に適しており、高次の学力を評価する際には、論述形式の課題を用いたり、知識や技能を総合的に活用することを求めるような複雑な課題を実施したりする必要がある。高次の学力の質的レベル（習熟の度合い）は、知識や技能を統合的に活用するようなパフォーマンスを要求する課題を通じて質的に評価される。

今日求められている学習・評価課題のデザインを端的に言えば、それは単一の絶対的な正答（一意に正しい応答）が存在せず、多様なアプローチが可能である課題といえる。たとえば、レポート、小論文、エッセイ、実験の考案と仮説の検証（実験設計）、オープンエンドな調査・分析、プレゼンテーション発表、創造的な作品制作、芸術パフォーマンスといった形式がこれに当たる。このような形式の課題は、発散的で開かれた応答を学習者に要求するため、「発散的課題（divergent task）」と呼ばれる²。発散的課題は、学習者に知識や技能の統合、複雑な問題解決、批判的・独創的思考などのパフォーマンスを披露する機会を提供す

る。学習者には、デザインとプロダクションの両方のコンピテンスが要求され、また、創造性を発揮するかなりの余地が与えられる。質の高いパフォーマンスを生み出す、単一的手段（一連のレシピやテクニック）は存在せず、また、単一の正答や解決策もない。

発散的課題による学習評価を「パフォーマンス評価」とも呼ぶ。松下は、パフォーマンス評価を、「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」と定義している³。

発散的課題の学習評価（パフォーマンス評価）では、特に、課題成果物をどう公平、公正に評価できるのかが、教師にとっては切実な問題となる。なぜなら、発散的課題は、単一的手段と正答が存在しないために、単純に正誤で評価できないからである。生徒のパフォーマンス（作品や実演）の質を判断する際には、多様な[多重の]（multiple）クライテリアを使用した熟達した質的判断が必要となる⁴。したがって、発散的課題をデザインし、その成果作品を評価することは、客観テストの場合と大きく異なる力量を教師に要求する。

では、(1)どのように学習・評価課題をデザインすれば、実りある学びを導き高次の学力を保障できるか、(2)どのように評価すれば、独善的、主観的な評価ではなく公平、公正な評価だといえるか、(3)課題デザインと評価に関する教師の力量は、どう形成されるか。

本稿は、以上の問いを出発点とし、発散的課題の学習評価における教師の力量形成に焦点を合わせる。本稿の目的は、(1)発散的課題による学習評価で教師に求められる力量がどのようなものか、その内実を詳らかにすること、(2)こうした教師の力量はどのように形成されるのか、そして、その形成をどのように助けることができるかを考察することである⁵。

第1章 教師に求められる評価力量の内実

本章では、発散的課題の学習評価において、教師に求められる力量の内実がどのようなものか明確にする。

第1節 課題デザインにおいて求められる理解と技術

学習・評価課題のデザインは、指導と学習の出発点として重要である。なぜなら、どう課題をデザインするかということが、一連の指導と学習の形式と展開を規定するからである。ゆえに、良い教育と良い学びは良い課題デザインから始まるといっても過言ではない。課題デザインを熟慮することは、どのような学力（資質・能力）を育むかを吟味することと同義である。

一般的に、指導や学習の過程に先立って、教育目標を明述し、出口における学習者の姿のイメージを膨らませることが大切である。ただし、それだけでは十分ではない。なぜなら、どれだけ高次の学力に対するコミットメントを目標として叙述しても、最終的にどのような課題に取り組み、課題でどの程度の出来栄を要求するのかを一定明確にしない限りは、目標叙述は絵に描いた餅になってしまうからである。この意味で、評価課題は、より実質的なカリキュラムを表現する⁶。

学習評価に関する多くの出版物では、評価課題の形式が「どのような学力を評価するか」ということを規定すると指摘されている。もちろん、両者の間には密接な関連がある。ただし、必ずしも課題形式の選択が、評価される学力の種類を決定するとは厳密にはいえない。発散的課題、および、客観的テスト形式のどちらも、高次と低次の学力を評価することができる。たとえば、論文やエッセイ形式の課題であったとしても、単純な知識の羅列や吐き出しにしかならないような場合もあり得る。よって、教師は、課題形式のみならず、「課題設定（課題文）」をどのようにデザイン（叙述）すべきか熟慮することが大切である。では、具体的に教師は、どう課題設定をデザインする必要があるのか。

通常、課題設定は教師によって定められるか、教師と学習者の共同作業によって定められる。いずれの場合においても、課題設定は、学習者が成果物として構築することを期待するパフォーマンス（作品や実演）のデザイン・パラメータを設定する必要がある。すなわち、どのようなパフォーマンスの発揮を求めるのか、その構成、形態、様相を、学習者に明確に伝える必要

がある。したがって、課題設定における文章の叙述だけでなく、いくつかの作品事例を提示することも大切である。課題をデザインする際、教師は、どのような学習者の応答を求めるのか明確にしていない、もしくは、殆ど考慮していない場合がある。しかし、教育を意図的な試みと捉えるのであれば、課題設定をデザインする際に、どのようなコンピテンスの育成を意図しているか、また、具体的にどのようなパフォーマンスの発揮を求めるのか一定明確にしておく必要がある。

教師がしばしば陥りがちな過ちは、たとえば、エッセイの課題において、どのようなトピック内容について論じるべきかを指定するが、どのような認知操作を要求するかは不明瞭というものである。たとえば、小論文において、「Aについて議論せよ」と単に内容のみを指定する課題設定は、目標や意図と照らし合わせて不十分な場合が多い。他方、「Aに関する言説Bと言説Cの相対的な意義と課題を分析して同定せよ」という設定は、批評を要求しているという点で、目指される小論文の「ジャンル（genre）」が明快である。小論文において、説明を求めるのか、批評や分析を求めるのかでは、学習者に要求される認知操作が異なる。重要なことは、課題設定が、明示的か否かは問わず、作品のジャンルを指定している必要がある点である⁷。

ジャンルの内で評価するという考え方は日常生活の中でも自然で一般的なものである。ジャンルの設定は、評価で正当に用いられるクライテリア群——ただし、明示的な場合もあれば、暗黙的な場合もある——を慣習的に規定する。評価は、価値についての判断である。ある評価対象が「素晴らしい質」と評価されるに値するならば、それは評価に用いられるクライテリア群がそのジャンル（文脈）の中で良いものとは何かの適切な定義として認められており、ある評価対象がそれらのクライテリアにおいて一定のパフォーマンス（出来栄）を示しているからである⁸。

たとえば、衣類の評価においては、その衣類がフォーマルな正装、スポーツウェア、寝巻きなどの種類として評価されるのかが、まず問われる必要がある。なぜなら、それぞれの種類の衣類には独自の目的があり、よって独自のクライテリア群への参照が必要になるからである。これは、小論文において、説明、分類、分析のどれを要求するかによって、異なる構成形式やレ

トリックが求められるのと同様である。課題設定において、作品デザインと評価のパラメータが定められることで、学習者のパフォーマンスが課題設定で定められた目的をどの程度達成しているか（提示された問題をどの程度うまく解決できているか）の評価が可能となる。つまり、課題設定は、価値判断の枠組みを規定する「基盤となる単一のクライテリア」として機能する。このように課題設定を的確に定めることは、高次の学力を保障する第一歩となる（妥当性の視点）。

以上の議論は、課題設定の出発点であるべきだが、課題設定で熟慮すべきことは他にもある。たとえば、(1)課題が学習者の知的関心や意欲を喚起させる魅力的なものとなっているか、(2)課題に取り組む有意さや切実さを学習者が感じるものとなっているか、(3)学習者にとってアクセシブルなものとなっているか、これらも重要な考慮事項である。教師には、以上を熟慮して、課題設定をデザインする力量が求められる。

第2節 質的判断において求められる理解と技術

では、課題設定に対する応答としての、学習者のパフォーマンス（作品や実演）は、どのように評価すべきか。その際、教師に求められる力量はどのようなものか。通常、熟達した評価力量を有する教師は、学習者のパフォーマンスに対する豊富な評価経験を有している。熟達した教師の知は、多くの評価経験——多様な作品事例（様々な質、多種多様なアプローチ）に晒されること——によって構成される。ただし、教師が学習者のパフォーマンスの中から優れたものを認識することが容易でも、必ずしも教師がそれを正確に描写できる訳ではない。教師が有する質の概念は、頭の中で明文化されない形態で保持されている場合が多い。

(1) 質の2次元空間を洗練させ拡張させる

教師は、多くの様々な学習者の作品事例に触れる経験の中で、質の2次元空間を豊かなものにする。この2次元空間の垂直軸は、質が高いか低いかという理解を表し、水平軸は、同等の質と判断される作品であっても様々な表現やアプローチの可能性が存在するという理解を表す。この2次元空間は、原則、過去の経験の範疇に依存するが、多くの経験を経ることで、2次元空間が濃密になると同時に、垂直方向と水平方向の領域が拡張される。力量ある教師は、学習者の作品を

2次元空間内の固有の位置に配置することができる。

ここでいう質について、教師が理解すべき重要な特徴がある。第一に、質は、単一の作品によって規定されるものではなく、複数の作品事例の背後に現れるという点である。同じジャンルの中には、優秀とみなされる多種多様なものがある。質は、異なる表面的な特徴を持っているものの、同程度の質と判断される作品事例群の背後に立ち現れるものである。

第二に、質とは、意図した目的を達成する為に、全体として作品（パフォーマンス）がどう構成されているのか、その出来栄を意味するという点である。ここで重要なことは、作品が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示していることである。ひとまとまりの質では、諸要素の組み合わせ、配置、関係が重要な意味を有するため、諸要素に単純に分割、合算したりすることはできない。分割、合算するのではなく、判断したい次元に即して直接的に質は認識される必要がある。

たとえば、「うどん」というジャンルの食べ物の質を評価するとき、「出汁」「麺」「トッピング」というクライテリア（諸要素）に言及することができる。ただし、「うどん」の質（出来栄）は、それぞれのクライテリア（部分的な質）に意識を向けながらも、「うどん」が全体としてどう構成され、どの程度の出来栄であるかという点において認識される必要がある。たとえば、「出汁」「麺」「トッピング」などの要素は、それぞれ最良であっても、それらが組み合わせられて一つの「うどん」になった時、それが最良でない場合がある。他方、あくまで「麺」の質を評価したいのであれば、たとえば「コシ」「細さ」というクライテリアに意識を向けながら「麺」が全体としてどう構成されていて、どの程度の出来栄であるかを判断する必要がある。

この段階で、スタンダードという概念を導入しよう。スタンダードとは、学習者のパフォーマンス（作品や実演）がどの程度優れているか、達成の度合いやレベルの固定点を指す。すなわち、先に論じた質の2次元空間における垂直軸の固定点がスタンダードである。相対評価ではない学習評価を行う際には、ある学習者のパフォーマンスが、どのスタンダードを満たしているかを判断することが必要になる。なお、単一のスタンダードを設定して評価する場合もあれば——この場

合、スタンダードは合格する上での最低ラインを示すことになる——A、B、Cと複数段階のスタンダードを設定し、どのスタンダードを満たすのかという評価を行う場合もある。

(2) 多様なクライテリアに精通する

学習者の作品に対して熟達した質的判断を行う——2次元空間内に配置する——ために、教師は、あるジャンルの評価において用いられる様々なクライテリアに精通する必要がある。ここでいうクライテリアとは、学習者のパフォーマンス（作品や実演）を評価する際に採用される観点、側面、枠組みを指す。クライテリアは様々な抽象度の次元を有するが、本稿では(1)評価枠組みを規定する価値体系としてのクライテリア（枠組みとしての価値軸）と、(2)質的判断において参照されるクライテリア（参照次元としての価値軸）を、便宜上区別したい。課題設定（課題文）が、価値判断の枠組みを規定するものとしての「基盤となる単一のクライテリア」になると先に指摘した。この意味で、課題設定は、評価枠組みを規定するクライテリアとして機能する。課題設定の枠組みに則ることで、そこで定められた目的を作品全体がどの程度達成しているかという、質（出来栄）の判断が可能になる。他方、作品の質を判断する際には、作品の様々な側面や観点到意識が向けられることになる。これらが、参照次元としてのクライテリアである。以後、クライテリアと単に言及する際には、後者の意味で使用する。

小論文を評価する際には、「論理性」「説得性」「わかりやすさ」「美しさ」「語彙」「首尾一貫性」「レトリック」「比喩」「パラグラフ構成」など多様なクライテリアが参照されて、作品の質が吟味される。作品の質を評価することに自信を持ち、評価に関するディスコースに関与するためには、教師は、様々なクライテリア概念に精通する必要がある。なお、異なるジャンルは異なるクライテリア群への参照を要求する。また、同一の用語であっても、その意味がどのジャンルにおいて用いられるかで変化することが多い。

小論文では「首尾一貫性」「わかりやすさ」「説得性」等のクライテリアがしばしば参照されるが、こうしたクライテリアは、単純に正・誤で判断できず、どの程度優れているのかは連続的に現れる（ファジーなクライテリア）。また、ファジーなクライテリアは、言語で

説明することができても、言語で明確に定義できない。したがって、これらクライテリアの意味に精通するためには、実在物である作品事例との対話を通して意味を掴み取っていく必要がある。

熟達した質的判断を行うために、教師は、このような複雑で多様なクライテリアの世界を自在に飼い慣らすことができなければならない。また、ある作品の質を見定めるときに、どのクライテリアに焦点を合わせるべきか、作品の様々な特徴や側面（クライテリア）がどのように絡み合っており、ひとまとまりの質をして構成されているかを掴み取れなければならない。

豊かな質の2次元空間を有し、多様なクライテリアに精通することで、教師は優れた質的判断を下すことが可能になる。こうした力量を備えることで、作品全体の質に対する視点と、ある特定のクライテリア（部分的な質）への局所的な視点を自由に切り替えながら、作品の質の程度（どのスタンダードを満たすか）を吟味していくことができるようになる。

(3) 間主観的な質的判断を下す

これまで論じてきた熟達した質的判断を下す力量は、教師が備えるべき必要条件ではあるが、十分条件ではない。特に、発散的課題の学習評価が総括的な目的で行われる場合、学習者のパフォーマンスが「どのスタンダードを満たすのか」という教師の判断を、十分に比較可能なものとする必要がある。すなわち、間主観的な質的判断を下す力量が求められるのである。

基本的に、評価とは価値の判断であるために、評価の有する主観性から逃れることはできない。マーク式の客観テストであっても、採点過程に人間の主観は存在しないが、作問過程において人の価値判断が反映されるために、それが「客観」テストを標榜していたとしても必ずしも客観的ではない。ただし、このように評価が人の主観から逃れられないものであっても、評価における間主観性を達成することは可能である。ある評価が、間主観的であるとは、複数人の評価者の間で評価判断の一致と同意が成り立っていることを指す。評価者間での間主観性の達成を目指すことで、評価の「信頼性（reliability）」の問題は、「比較可能性（comparability）」の問題として再定義される。

このように教師の質的判断を調整する過程をキャリブレーションと呼ぶ。その語源 *calibrate* は、計器の

目盛りを正しく調整することを意味する。つまり、評価が行われる前に、計器や人間の判断が統一されるように正しく調整することをキャリブレーションと呼ぶ。言い換えれば、評価の比較可能性を達成する上で、評価者の「物差し」を調整する過程ともいえる。なお、どの程度の範囲（たとえば学校内や学校間）で、比較可能性を達成するかという点も念頭に置く必要がある。一般的に、比較可能性を達成しようとするコミュニティの範囲が狭いほどそれは容易で、範囲が広がることで困難さが増す。教師には、要求される範囲内で比較可能な判断を下すことが求められる。

以上、本章では、発散的課題による学習評価で教師に求められる力量の内実について詳らかにしてきた。

第2章 発散的課題における教師の評価力量の形成

本章では、以上で論じた評価に関する教師の専門的力量がどのようにして形成されるか、その形成をどう助けることができるかを論究する。これまでの議論を踏まえると、これは、(1)いかにして教師の質的判断を洗練できるか、(2)比較可能性（間主観性）が達成できるかという2つの問いに帰着できる。これらの問いに対して、今日、最も幅広く支持され受容されている手段は、「ルーブリック」という装置を用いることである。本章では、特に、このルーブリックなるものに焦点を合わせて、その可能性と危うさについて検討する。

第1節 ルーブリックとは何なのか

近年、ルーブリックは学習評価を語る上で欠くことのできないキーワードである。ルーブリックを旗印に学習評価改革を進める国際的動向は、一まとまりの意味システムとして強力な磁場を形成している。ルーブリックの導入によって、教師の質的判断による評価（発散的課題による学習評価）を重視し推進しつつも、そこでの評価が独善的で恣意的なものに陥ることなく、透明性や客観性を一定担保することが目指されている。

ただし、ルーブリックといっても、ルーブリックと称されるものの内実（目的・形態・使用法等）は想像を超えるほどに多様である。唯一共通することは、それらがクライテリアとスタンダードに関する言語表現マトリックスということである。たとえば、総括的評価の視点では、それは、採点装置、キャリブレーション

の装置（間主観性を達成する装置）、教師の質的判断を洗練させる装置、アカウントビリティの装置などとして機能することを目的に導入される。他方、形成的評価の視点では、学習者のピア評価・自己評価を助ける装置、フィードバックやフィードフォワードの装置等として機能することを目的に導入される。

本稿の関心は、教師の力量形成にある。したがって、本稿では、(1)ルーブリックを用いることで教師の質的判断を洗練させることができるのか、(2)ルーブリックを用いることで教師の質的判断を比較可能なものにするのかという点に限定して、ルーブリックの可能性と危うさについて検討する。なお、前者は後者の必要条件であるが、十分条件でないことに注意されたい。

本稿の基本的立場を先取りすれば、ルーブリックは評価枠組みを外化し、間主観的な質的判断への可能性を目指す点にその意義が認められるが、実際にその目的を達成できているか疑問が残るということである。また、特に実践段階においては、その形態や運用方法に多くの問題点があるということである。ただし、ルーブリックが目指すこと、それ自体は追求すべきことであり、ルーブリックが孕む諸問題を意識した上で、ルーブリックを適切に飼い慣らすことを目指すか、根本的に異なるアプローチを模索することが重要となる。

第2節 ルーブリックの可能性と危うさ

前章で論じた通り、評価に熟達している教師は、豊かな質の2次元空間を頭の中に所有している。ルーブリックは、この2次元空間を構成する評価熟達知を、言語表象として外化を試みるものであるといえる。質の固定レベルとしてのスタンダードを明示し、どのようなクライテリアが評価に用いられるのか言語で叙述するのである。ある質の水準にある作品は、それぞれの表面的特徴が異なりつつも、その背後に共通的部分が存在する。この共通的な性質を言語で叙述しようと試みるのである。なお、これらは、全体的に構成される場合もあれば、観点ごとに叙述される場合もある。

このように評価枠組み（スタンダードとクライテリア）を明示することは重要である。なぜなら、評価枠組みを外的に形成することは、教師の質的判断の比較可能性を達成する第一歩となるからである。評価枠組みが外化されることで、質的判断はそれぞれの教師の

個人的で主観的なものでなく、評価枠組みに準拠したものとなる。また、評価枠組みが外化されることで、教師のみならず、学習者やその他利害関係者にもそれらを公開することができる（透明性の達成）。

明示的な評価枠組み、すなわち、ルーブリックが外的に構成されることで、各教師がそれに準拠した質的判断を行うことが可能になる。つまり、ルーブリックはキャリブレーションの装置として機能する。また、ルーブリックを用いることで、経験がまだ浅い（十分な質の2次元空間を有していない）教師の質的判断をさらに向上させることが期待されている。ただし、それほど単純ではないというのが本稿の主張である。

(1) ルーブリックによる比較可能性達成の問題点

ファジーなクライテリアが重要な位置を占める学習評価においては、言語記述そのみでスタンダードを固定することはできず、結果として、スタンダードを適切にキャリブレートすることができない。なぜなら、スタンダードの記述文は直感的に理解できるように思えるが、そこで用いられる語句や修飾詞は、具体的な文脈無くして解釈することはできず、決定的意味を持つことはないからである。また、スタンダードの記述では、相対的な修飾語を用いるために、スタンダード記述の修飾語は、絶対的ではなく相対的に解釈される。たとえば、ルーブリックの場合、全く同じルーブリックが、学部生、大学院生でも同様に用いることができる場合が多く（求められる水準は大きく異なるのにも関わらず）、このことは、ルーブリックがクライテリアは指定しても、スタンダード（求められる質の水準）は指定できていないことを示唆する。よって、教師集団でルーブリックの記述に対する同一の解釈を達成するためには、必ず複数の作品事例群を参照する必要がある。また、作品事例群とルーブリックを関連づける共通の評価ディスコースを形成する必要がある。

ただし、作品事例群とルーブリックを直接的に紐付けることは困難である。なぜなら、ルーブリックの記述は、あるスタンダードの一般的な特徴を叙述したものであり、あるスタンダードに共通的に出現するクライテリアに焦点化されているからである。よって、ルーブリックは、同一のクライテリア群を全ての作品に画一的に適用する実践を引き起こす可能性がある。ただし、共通に出現するクライテリアが、必ずしも各作

品の評価判断において最も重要なクライテリアではない。たとえば、「最高」と評価される作品群に共通に出現するクライテリアと、「最高」と評価される個々の作品の評価で最も重要なクライテリアは異なる。ゆえに、どのようなクライテリアがどのような作品に適用されるのかという、画一的ではない共通理解が求められるが、ルーブリックは、この種の情報を含まない。

(2) ルーブリックで評価知を形成することの危うさ

ルーブリックは、あくまで共通の性質を明文化するものであるため、それを作る元となった複雑な評価知識の一部は捨象される。ルーブリックは質的判断の複雑さを完全に表現できない。そのため、ルーブリックの記述をどれだけ読んだところで、背後にある評価知識がどのようなものか掴み取ることはできない。洗練された質的判断を行うためには、ルーブリックが作られる元となった複雑な評価知識（質の2次元空間の知）を構成しなくてはならないのである。

力量ある教師は、深い評価知識とそれらの複雑な運用規則を知っていることで、時にルーブリックの記述を破りながら、ルーブリックには記述されていない背後の意図も踏まえ、ルーブリックを飼い慣らすことができる。ただ、そうでない者がルーブリックそれ自体の記述によって複雑な評価知識を形成すること（質的判断の技芸の本質を掴むこと）はできず、場合によっては、絶対的な規範に見えるルーブリックを手にすることで、それが足枷ともなる。場合によっては、ルーブリックの記述をただ機械的に当てはめていくといったような実践を引き起こす可能性がある。この場合、ルーブリックは、単純な採点プロトコルに成り下がる。

質的判断という複雑な営みを操るためには、明文化された価値・規則体系を適用する技量の形成ではなく、事例の中に内在する価値・規則体系をある種暗黙的に掴み取ることで、複雑な評価知識を形成する必要がある。理想的なルーブリックとは、これを型の形式で外化するものであるといえる。よって、熟達した質的判断の技量をすでに備える教師が、参照とする枠組みとしては一定有用に機能しうる。ただ、複雑な評価知識を、ルーブリックそれ自体によって単純に共有化することはできない。複雑な評価知識からルーブリックを作り出すことができても、ルーブリックそれ自体から複雑な評価知識を作り出すことはできないのである。

作品事例における質をホリスティックに掴み取る行為と、ルーブリック等で明述されたクライテリアを当てはめて評価していく行為は、根本的に異なる認知操作を要求する。部分的な質（各クライテリアにおける質）を手がかりとしつつも、作品全体の質を掴み取り、その論拠をクライテリアに言及しながら正当化するというプロセスによって、質の2次元空間は、徐々に豊かなものになっていくのである。

(3) 具体的な作品事例からスタートする

質の2次元空間を豊かなものにするためには、多くの作品事例に晒される必要があり、それらの全体的な（一まとまりの）質を掴み取る経験が必要となる。この際、教師集団において共通の評価ディスコースを形成する手段として、ルーブリックや明示的なクライテリア群を参考にするには、一定の有用性があるだろう。それらを使いこなす術を知っている先輩教師のもとで、具体的な事例に対する評価経験を共有することで、言語記述が意味することを具体的で明瞭なものとし、ルーブリックの記述を生きたものとしていくのである。

したがって、教師集団が共同して、生徒の具体的な作品事例と対話して、ルーブリックを作成したり問い直ししたりする営みは、一定の価値ある実践である¹⁰。ルーブリックの作成過程に参画した教師らは、その過程で、お互いの質の2次元空間を共有し合い、共通の評価ディスコースを形成し、お互いの2次元空間を擦り合わせ、合意を形成することが可能となる。この過程において、教師らの質的判断が洗練されたり、キャリブレーションされたりすることとなる。

ただし、留意すべきことは、結果として生み出されたルーブリックの記述それ自体のみでは、質的判断を洗練させたり、キャリブレーションを行ったりする装置として有効に機能することはできないということである。繰り返すが、その理由は、ルーブリックの背後にある複雑な評価知識が欠落していることによる。この活動の利点は、あくまで、その活動に参加した教師のコミュニティ内でのみ発揮されるのである。

第3節 2つのアプローチ

ここで、本章の問い——(1)どう教師の質的判断を洗練させることができるのか、(2)どう教師の質的判断を比較可能（間主観的）なものにできるのか——に立ち

戻ろう。本稿では、次の2つの可能性を提示する。(1)留意すべき事項を押さえた上で、ルーブリックを飼い慣らす方法、(2)ルーブリックとは根本的に異なる方法で、その目的の達成を目指す方法の2つである。

(1) ルーブリックを飼い慣らす

これまで示してきた通り、ルーブリックは多くの危うさを孕んでいる。しかし、その危うさに教師が意識的に留意することで、ルーブリックを飼い慣らすアプローチを目指すことも一つである。その際、以下のことに留意する必要がある。

(1)ルーブリックは、厳密な評価プロトコルとして認識されるべきではない。ルーブリックを有効に用いるためには、ルーブリックを構成する元となったパフォーマンス（作品や実演）の事例群との対話が必要になる。ルーブリックは、具体的な事例群を共有した上で、各教師が有する質の2次元空間を擦り合わせる際に参照とする「やわらかい枠組み」として捉えるべきである。(2)ルーブリックにおける言語表現が意味することが何であるか、パフォーマンス（作品や実演）の事例群を共有した上で、専門家コミュニティで同一の評価ディスコースの形成に努める必要がある。(3)各クライテリアにおける（部分的な）質に焦点を向けながらも、それらがどのようにまとまり作品全体の質を形作っているか吟味しなくてはいけない。部分的な質は、全体的な質を達成すべき手段であって、それ自身が目的ではない。(4)各クライテリア（観点）の得点を合算して全体の質の水準を判断することはできない。あくまで、全体的な質については、それに対する直接的な判断が必要になる。各次元の質は次元ごとに判断する必要がある。(5)ルーブリックに示されるクライテリアは、あくまで質を判断する際の手がかりに留める必要がある。同一のクライテリア群を全ての作品において、画一的に適用するべきではない。共通的なクライテリアと評価に用いるべきクライテリアは異なるからである。(6)ルーブリックの記述は、常に不完全な表現であるために、常に破れる可能性があることに留意する。創成クライテリアが出現するなどして、ルーブリックには記述されていないクライテリアが出現する場合、これを考慮する必要がある。また、必要に応じてルーブリックの枠組みそれ自体を問い直し再構成しなければならない。

(2) 根本的なオルタナティブを模索する

他方、ルーブリックとは根本的に異なる方法でアプローチすることも可能である。たとえば、豪州の教育評価研究者ロイス・サドラーは、ルーブリックを用いずに、作品事例に基づいてスタンダードを外化することを主張している¹¹。サドラーはスタンダードの外的表現として、あるスタンダード・レベルを満たす作品事例の集合を構成し、それぞれの作品事例に、その判断に対する論拠説明が付すことを提案している。各々の作品の論拠説明では、同一のクライテリア群が画一的に適用されるのではなく、それぞれの作品を評価するのに最も適切なクライテリアが用いられて記述される。彼は、重要なことが、スタンダードを定義する出発点が言葉による叙述ではなく、具体物である作品事例にあることを強調する。ただ、作品事例の集合だけでは、それらの質をどのように見ていくのかを掴み取ることが難しく、それらを助けるガイドが必要となる。そこで、言語記述（論拠説明）が重要な役割を果たす。

ルーブリックとは根本的に異なるアプローチを目指す背景には、ルーブリックの次の特性が本質的に回避であるという認識があると考えられる。ルーブリックは、明示的に評価枠組みを提示するために、シンプルであり、わかりやすいものである。しかし、こうした特徴がときに暴走を引き起こす。ルーブリックの多くは、——それが必ず作品事例の添付を必要とするのにも関わらず——記述のみが一人歩きしてしまうことが多い。このように記述のみが一人歩きしたり、ルーブリックが厳格な採点指針として扱われたりする事態によって、ルーブリックの多くが評価プロトコルに陥ってしまっているという現状がある。

このような作品事例の集合によるアプローチは、可搬性や煩雑性という視点から、あまり現実的ではないと指摘されることがある。ただし、近年は、作品事例のデータを、PCやタブレット、クラウド等で保存、管理し共有化することが格段に容易になっている。したがって、ICT 技術を活用することで、作品事例を参照したり添付したりすることは困難なことではない。

おわりに

以上、本稿では、(1)発散的課題による学習評価で教師に求められる力量がどのようなものか、その内実を

詳らかにし、(2)こうした教師の力量がどのように形成されるかについての考察を行った。

最後に本稿での議論を総括しよう。第1章では、教師には、(1)目標を熟慮しながら適切な課題設定をデザインし、(2)課題設定を踏まえて、多様なクライテリアを自在に操りながら、生徒のパフォーマンス（一まとまりの質）の水準を判断し（質の2次元空間を豊かにし、多様なクライテリアに精通することでこれが可能となる）、(3)比較可能（間主観的）な質的判断を下す力量が求められることを指摘した。第2章では、ルーブリックを教師の比較可能性を達成する装置、質的判断を洗練させる装置として活用する際に生じる可能性と危うさを詳らかにし、(1)留意してルーブリックを飼いつづける、又は、(2)ルーブリックとは抜本的に異なる方法で、その目的の達成を目指すという2つアプローチを見出すことができると主張した。

注

- 1 石井英真『今求められる学力と学びとは』日本標準、2015年。
- 2 Sadler, D. R. (2009). Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughin (Ed.). *Assessment, learning and judgement in higher education*. 45-63. Springer.
- 3 松下佳代「パフォーマンス評価による学習の質の評価」『京都大学高等教育研究』18巻、2012年、pp.75-114。
- 4 Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- 5 本稿のアイデアの多くは、著者とロイス・サドラーが、2022年の1月から2月にかけて、豪州クイーンズランド大学で行った学術的談義に基づいている。
- 6 ロイス・サドラー、石田智敬（訳）「カリキュラムを表現するものとしての評価課題」『カリキュラム研究』29巻、pp.91-109。
- 7 Sadler, D. R. (2014). Learning from assessment events: The role of goal knowledge. In Kreber, C., Anderson, C., Entwistle, N., & J. McArthur (Eds.). *Advances and innovations in university assessment and feedback*. 152-172. Edinburgh University Press.
- 8 Sadler, D. R. (1985). The origins and functions of evaluative criteria. *Educational Theory*, 35, 285-297.
- 9 石井英真「レポート課題を評価するとき」成瀬尚志編『学生を思考にいざなうレポート課題』ひつじ書房、2016年。
- 10 西岡加名恵『教科と総合学習のカリキュラム設計』図書文化、2016年。
- 11 石田智敬「スタンダード準拠評価論の成立と新たな展開」『カリキュラム研究』30巻、2021年、pp.15-28. Sadler, D. R. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, 67(3), 273-288.

（日本学術振興会特別研究員・博士後期課程）

受理 2022年2月26日