

プログラムレベルと科目レベルの評価をつなぐ

— PEPAの理論と課題 —

松 下 佳 代
(京都大学)

Combining Course- and Program-Level Outcomes Assessments: Pivotal Embedded Performance Assessment's Theory and Its Challenges

Kayo Matsushita
(Kyoto University)

Pivotal Embedded Performance Assessment (PEPA) is one promising approach for the learning systems paradigm our project pursues. The purpose of this paper is to focus on PEPA from the viewpoint of combining course- and program-level outcomes assessments and to clarify its underlying ideas and challenges. Generally, program-level assessments include a sum-based method, such as GPA or a curriculum map, and a trajectory-based method, such as learning portfolios. “Sum” and “trajectory” represent the two aspects of a curriculum. PEPA can be thought as a mediation of these two methods. Its core feature is to grasp the trajectory of student learning by assessing the milestone learning outcomes in the key courses of the degree program and sequencing them toward the program goals. Our future challenges are the enlargement of the disciplines PEPA can be applied to and the improvement of its feasibility.

〔キーワード：プログラムレベルの評価, PEPA (重要科目での埋め込み型パフォーマンス評価), 総和と軌跡, カリキュラムと評価の整合性, PBL〕

1. 本課題研究とPEPAの位置づけ

現在の大学教育の焦点は、「教員が何を教えたか」から「学生が何を学び、身につけたか」に移ってきている。大学教員には、自分の担当科目だけでなく学位プログラム全体を通して、学生がどう学び成長しているかを把握することが求められる。だが、DP（ディプロマ・ポリシー）に記述される学位プログラム全体の学修成果は、抽象的で捉えどころのないものになりやすい。そこで本課題研究では「学習システム・パラダイム」への転換をめざす。それは「大学教員が学習者の視点に立って、担当する授業科目だけでなく、プログラムの全体性・整合性も重視する認識の枠組み」（本号深堀論文）のことである。

この取組において有望なアプローチの一つが、「重要科目での埋め込み型パフォーマンス評価 (Pivotal Embedded Performance Assessment: PEPA)」（Matsushita, Ono, & Saito, 2018; 松下, 2019a）である。PEPAでいう「重要科目」とは、その授業科目の目標がプログラム全体の目標に直結する科目のことである。それはしばしば、それまでに学んだ知識やスキルを統合し、高次の能力を育成・発揮することを求める科目でもある。「埋め込み型」とは、特定の科目の評価が同時にプログラム・機関レベルの評価の働きももつようにしたものであり、カリキュラムとは別に標準テストや学生調査などの形で行われる「追加型 (add-on)」と対比される (Suskie, 2009)。さらに、「パフォーマンス評価」とは、学習者のパフォーマンス（作品や実演など）を手がかりに、概念理解の深さや知識・スキルを統合的に活用する能力を評価する方法のことである。

本稿は、プログラムレベルと科目レベルの評価をどうつなぐかという問題意識に立って、PEPAを取り上げ、

その理論と課題を明らかにすることを目的とする。

2. プログラムレベルの評価方法—〈総和〉と〈軌跡〉—

(1) 従来的方法

プログラムレベルの評価として、従来どんな方法が用いられてきたのだろうか。毎年実施されている文部科学省の調査「大学における教育内容等の改革状況について」、および、朝日新聞・河合塾の調査「ひらく 日本の大学」によれば、主な方法として、標準テスト、学生調査、成績評価の活用（GPAなど）、卒業論文・卒業研究、学修ポートフォリオなどを挙げることができる（表1）。

ここに示すように、それぞれの評価方法には、メリットの一方で、限界もある。例えば、学生調査（学修行動調査など）は、朝日新聞・河合塾（2018）によると、全学実施の割合が2番目に高い方法となっているが（63%）、これ単独でプログラムレベルの学修成果の評価を行うのは不適切である。例えば、アメリカのアクセディテーション団体であるWSCUC（WASC Senior College and University Commission）のガイドブックでは次のように書かれている。「学生の学習のエビデンスは、調査（コンピテンスや成長についての学生による自己報告）以上のものを含むべきである。[中略] 学生の学習のエビデンスを提供するには、学生の学習の成果物の直接評価の結果を含む、より多くの異なるタイプのエビデンスが求められる」（WSCUC, 2015, p.16）。

(2) 〈総和〉による方法

では、直接評価の方法にはどんなものがあるだろうか。表1に挙げた方法のうち、量的／質的な直接評価であり、全学実施の割合も最も高い（92%）のは、「成績評価の活用（GPAなど）」である（朝日新聞・河合塾, 2018）。だが、各授業科目の成績評価を寄せ集めただけでは、プログラムレベルの評価にはならない。また、GPAはプログラム期間全体の成績評価の代表値ではあっても、DPに掲げられた複数の学修成果それぞれの

達成度を示すものではない。

そこで用いられるのが、カリキュラムマップやレーダーチャートである（Jankowski & Marshall, 2017；朝日新聞・河合塾, 2012）。カリキュラムマップは、プログラムレベルの学修成果が各科目によってどう達成されるかをマトリックス形式で示した表であり、レーダーチャートはそれを一つの図で表したものである。

カリキュラムマップの背後にあるのは、「個々の科目の学習成果の総和が、学習プログラム全体としての学習成果となる」（González & Wagenaar, 2008, 邦訳 p.33）というアイデアである。これを「〈総和〉による方法」と呼ぼう。また、レーダーチャートに示される各学修成果の達成度は普通、各科目の成績を何らかの計算式（加重平均など）を使って足し合わせることで示されるので、これも〈総和〉による方法に依拠しているといえる。

だが、〈総和〉による方法には次のような難点がある。

(a) カリキュラムマップで描かれるプログラムレベルの学修成果は、DPで記述されるようなかなか抽象度の高い資質・能力である。それらは多義的で曖昧であり、科目の内容によって、また担当教員の解釈によって異なる意味をもちうる。「全く異なるさまざまなスキルや能力の緩やかな集合体に割り当てられた名称にすぎない」（Kosslyn & Nelson, 2017, p.29）ことが多いのである。プログラムレベルの学修成果の意味が科目によって異なるのだとすれば、それを数値化して足し合わせることにどんな意味があるのだろうか。

(b) 〈総和〉による方法は、時間軸を捨象している。だが、多くの学修成果は、プログラムの期間内で大きく変化する。例えば、「論理的に文章を書く力」のような学修成果は、1年次の初年次教育科目と4年次の卒業論文のどちらにも含まれるが、卒業時に学生が身に付けている力は卒業論文の方でより直接に示されているはずである。だが、〈総和〉による方法では、こうした資質・能力の変化は考慮に入れられない。

表1 プログラムレベルの評価の方法

評価方法	タイプ	評価時期	メリット	限界
標準テスト (ジェネリックスキル)	・量的な直接評価	・通常、プログラム期間に2回以上	・組織間比較や経年比較が行いやすい ・教員の評価負担が小さい	・DPと合致しているとは限らない ・分野固有の知識や能力は測定できない ・ペーパーテストで測定できる範囲に限られる
学生調査 (学修行動調査など)	・量的な間接評価	・通常、プログラム期間に2回以上	・組織間比較や経年比較が行いやすい ・教員の評価負担が小さい ・調査できる範囲が広い	・自己報告のため、直接評価の代替はできない
成績評価の活用 (GPAなど)	・量的／質的な直接評価	・プログラム期間全体	・成績評価の結果が直接活用できる ・すべての科目がカバーできる	・各科目での学修成果の質の違いが捨象される ・総和では、学生の学びの軌跡や卒業時点での学修成果が把握できない
卒業論文・卒業研究	・質的な直接評価	・4年次	・単なる評価課題ではなく学習課題としての意義をもつ ・卒業時点での統合的な能力が把握できる	・4年次まで把握できない ・評価が主観的になりやすい
学修ポートフォリオ	・質的な直接評価 (根拠資料による)	・プログラム期間全体	・根拠資料とともに学生の学びの軌跡が把握できる	・カンファレンス（他者とのリフレクション）や学修成果との対応づけが伴わないと、根拠資料を保存しただけになる

(3) 〈軌跡〉による方法

プログラムレベルと科目レベルの評価をつなぐ方法には、〈総和〉による方法以外にもう一つ別の方法がある。それが、〈軌跡〉による方法である。カリキュラム研究では、カリキュラムという概念を「学習者に与えられる学習経験の総体」とも、また「学びの経験の履歴」とも捉えるが（佐藤，1996），〈総和〉による方法は前者に、〈軌跡〉による方法は後者に対応するといえる。

「卒業論文・卒業研究」は、〈軌跡〉による方法の性格を内包している。卒業論文・卒業研究は、学びの軌跡を経て卒業時までを得た資質・能力を総合的に反映していると考えられるからだ。また、「学修ポートフォリオ」も、根拠資料とともに学生の学びの軌跡が把握できる点で、まさしく〈軌跡〉による方法だといえる。だが、「卒業論文・卒業研究」は4年次まで把握できず、評価が主観的になりやすい。また、「学修ポートフォリオ」は、多様な根拠資料を含むため、目標とする学修成果との対応づけが適切になされることが少なくない。

私の知る限り、〈軌跡〉による方法を最も徹底的に追求しているのは、ミネルヴァ大学である（Kosslyn & Nelson, 2017；松下，2019b）。ミネルヴァ大学は、世界の7都市を移動しながら大学生活を送ること、すべての授業をオンラインのアクティブラーニングで実施することなどで知られているが、本稿の目的に照らして注目されるのは、カリキュラムの体系化と評価の方法である。

ミネルヴァでは、抽象的で多義的・曖昧になりがちな汎用的能力を100近い「知の習慣（Habits of mind）」と「基本的概念（Foundational concepts）」（まとめてHCsと呼ばれる）に分析し、4年間の正課科目や準正課活動を通じて習得できるよう、カリキュラムが体系化されている。1年次でのHCsの成績は、その後もさまざまな科目や正課・準正課のプロジェクトの中でHCsが活用されるたびにたえず更新され、卒業時の成績が学士課程全体を通じて育成された学生の汎用的能力の習熟度を示すものとなる（「タイムトラベル”グレード（“time-traveling” grades）」と呼ばれている）。

ミネルヴァ大学のようなやり方は、日本のほとんどの大学では実施できないだろう。だが、その評価の方法は、〈軌跡〉による方法の究極的な形を具体的に示したものとして興味深い。

3. PEPAの性格と手続き

(1) PEPAの性格

PEPAは、プログラムレベルと科目レベルの学修成果の評価をつなぐシステムとして、新潟大学歯学部での取

組をふまえて開発された方法である（Matsushita et al., 2018）。第2章で述べてきたことと関連づけていえば、〈総和〉による方法と〈軌跡〉による方法の中間型といえることができる。

新潟大学歯学部では、PEPA以前にも、プログラムレベルの評価として、臨床実習前の共用試験（CBT, OSCE）や国家試験があり、卒業時アンケートも実施されていた。また、新潟大学全体としては、カリキュラムマップにもとづく評価（NBAS）も行われている。だが、知識・スキルを統合的に活用する能力をみる質的な直接評価や、学生の学びの軌跡を把握する評価は不十分であった。PEPAはこうした点を補完する意味をもっている。

(2) PEPAの手続き

PEPAは、学習システム・パラダイムの特徴であるカリキュラムと評価の整合性（alignment）を明確に意識したアプローチである。

新潟大学歯学部（歯学教育プログラム）のDPでは、プログラムの到達目標が、「知識・理解」「専門的能力（分野固有の能力）」「汎用的能力」「態度・姿勢」という大学共通の4つの柱にそって、24項目設定されている。さらに、CP（カリキュラム・ポリシー）では、「本プログラムでもっとも重視する学習成果である歯科臨床能力は、歯科医療という文脈における問題解決能力と定義できる。低学年から高学年に向けて、問題解決能力から歯科臨床能力へと専門性・総合性・真正性を高めて育成し、その学習成果をプログラムの教育目標に直結する重要科目で直接評価して卒業生の質を担保する」と述べられている。この中の「重要科目での直接評価」がPEPAにあたる。

新潟大学歯学部におけるPEPAは、以下の手続きで構成されている。

①カリキュラムの体系化・分節化

学位プログラムを体系化・分節化し、プログラムの目標と各科目の関係を明らかにする（歯学教育プログラムは大きく4期に分節化）。

②重要科目の抽出とパフォーマンス評価の開発・実施

各期の授業科目から重要科目を1つ選び、そこでは教員団でパフォーマンス評価を開発・実施する（大学学習法、PBL、模型・シミュレーション実習、診療参加型臨床実習の4つを重要科目に設定）。それ以外の個々の科目の評価は、担当教員のエキスパート・ジャッジメントに委ねる。

③パフォーマンス評価における合格基準の設定と形成的評価としての機能

ルーブリックの各観点で一定レベル以上を合格とす

る（現状では、合格基準を「レベル1以上」、望まれる到達度を「レベル2以上」に設定）。合格基準が達成できるよう科目をシリーズで配置する（「大学学習法1・2」「PBL1・2」など）。

④プログラムの修了認定

重要科目での合格基準の達成と所定の単位の修得をもって修了認定する。

重要科目でのパフォーマンス評価については詳しく紹介することができないが、既にさまざまな形で公表されているので、そちらを参照していただきたい（「大学学習法」は、小野・松下，2016；松下，2019a など。「PBL」は、小野他，2014；小野・松下，2015；小野他，2018など。「模型・シミュレーション実習」は、秋葉他，2017など。「診療参加型臨床実習」は、藤井他，2017；小田他，2017など）。

(3) PEPAの中核的要素

上述のように、新潟大学歯学部におけるPEPAの取組は、教員団と共同研究者による8年あまりの実践と研究を経て構築されてきたものであり、他の大学・学部が取り入れるのはそれほど容易なことではない。だが、その中核的要素は比較的シンプルである。すなわち、学位プログラムの節目に配置された重要科目において、学生の中間的な学修成果を直接評価し、それを系列化することで、学生の学びの軌跡を把握するということである。重要科目は、たいていの場合、それまでに学んだ知識やスキルを統合し、高次の能力を育成・発揮することを求める科目であるので、その科目だけでなく、他の科目での学修成果も一定程度、把握できることになる。また、いうまでもなく、重要科目以外の科目は別途、評価が行われる。歯学教育の場合は、知識や認知的スキルだけでなく、社会的スキル・身体的スキルなども要求されるので、重要科目におけるパフォーマンス評価はかなり手の込んだものになっているが、他の分野では、プロジェクトやレポート、プレゼンテーションなどの形で実施することができるだろう。

2020年1月に、中教審大学分科会から「教学マネジメント指針」が公表されたが、上述のようなPEPAの中核的要素は、その中でも「学修成果の可視化」の方法として言及されている¹。

4. PEPAの課題

以上では、わが国で従来用いられてきたプログラムレベルの評価方法について検討し、プログラムレベルと科目レベルの評価をつなぐ際に、〈総和〉による方法と〈軌跡〉による方法があることを指摘した。その上で、

PEPAが両者の中間型であり、〈総和〉による方法の不十分さを補うものであることを示してきた。最後に、PEPAの課題として、「分野の拡大」と「実行可能性の向上」について述べておきたい。

(1) 分野の拡大

これまでPEPAに取り組んできたのは新潟大学歯学部や医療系単科大学の藍野大学だが、PEPAは医療系だけでなく幅広い分野に適用可能なアプローチである。

現在、東京都市大学では、PBLの先進校であるオーホルボー大学（Aalborg University）のモデルを取り入れながら、全学でPEPAに類する取組が始まっている。東京都市大学の場合は、独自に立ち上げた「SD PBL（Sustainable Development Project organized Problem Based Learning）」をPEPAでいう「重要科目」として設定し、初年次教育から卒業研究へとつながる学びの履歴を作ろうとしている。「SD PBL」はその名の通り、持続的社會への変革を志向するプロジェクト型のPBLである。1年次のSD PBL(1)では学科内のチームで汎用的リサーチメソッドを身につけ、2年次のSD PBL(2)では学科間のチームで専門のリサーチメソッドを習得し、3年次のSD PBL(3)では学部間のチームで専門以外のリサーチメソッドを知って今までの学びを俯瞰・体系化し、その上で卒業研究につなぐ。このように、(1)、(2)、(3)と進むにつれて、総合性・真正性が高まり、チーム・メンバーの多様性が増していく。新潟大学歯学部の場合は、学年が上がるにつれて専門性を高次化させていく方向をとっているが、東京都市大学の場合は、他の専門分野も学ぶことで俯瞰性を高めようとしているのが特徴的である。

今後、東京都市大学の事例も含め、さらに適用分野を拡大しながら、PEPAの理論をより深めていきたい。

(2) 実行可能性の向上

PEPAでは、評価負担の大きいパフォーマンス評価を重要科目に限定し、他の科目は個々の教員のエキスパート・ジャッジメントに委ねることで、評価の実行可能性を担保しようとしている。重要科目の評価を協働で開発・実施することで、各教員のエキスパート・ジャッジメントが涵養される。また、重要科目の評価自体も、次第に評価負担が軽減されていく（例えば、PBLの評価は、当初すべての学生について3名の教員で行っていたが、評価者間信頼性が確認されたことから、今では1クラスを2名で分担して実施している）。

このように、PEPAの取組は実行可能性を高める上でも有効であると考えているが、これはまだ仮説にとどまっている。今後は、本課題研究で開発された指標・尺

度なども用いながら、ケーススタディを通してこの仮説を検討していきたい。

注

¹ 「教学マネジメントを確立する上で各大学の判断の下で収集することが想定される情報の例」として、「『卒業認定・学位授与の方針』に定められた特定の資質・能力の修得状況を直接的に評価することができる授業科目における到達目標の達成状況」が取り上げられている（中教審，2020，p.26）。

文献

- 秋葉奈美・長澤麻沙子・小野和宏・前田健康・魚島勝美（2017）「新潟大学歯学部における統合型模型実習の取り組み」『日本歯科医学教育学会雑誌』33，106-114.
- 朝日新聞・河合塾（2012）「『ひらく 日本の大学』第4回：学修成果の把握」Guideline，2012. 11，25-35. (https://www.keinet.ne.jp/gl/12/11/hiraku_1211.pdf)（2019年11月12日アクセス）
- 朝日新聞・河合塾（2018）「学修成果の把握・可視化—ひらく 日本の大学調査より—」Guideline，2018. 11，71-79. (<https://www.keinet.ne.jp/gl/18/11/seicho.pdf>)（2019年11月12日アクセス）
- 中央教育審議会大学分科会（2020）「教学マネジメント指針」2020年1月22日. (https://www.mext.go.jp/content/20200206-mxt_daigakuc03-000004749_001r.pdf)（2020年2月8日アクセス）
- 藤井規孝・竹中彰治・多田田康一・佐藤直子・秋葉奈美・小田陽平・勝見祐二・小野和宏・前田健康（2017）「新潟大学歯学部臨床実習における臨床能力評価」『日本歯科医学教育学会雑誌』33，4-11.
- González, J., & Wagenaar, R. (Eds.). (2008). *Tuning Educational Structures in Europe, Universities' contribution to the Bologna Process: An introduction* (2nd ed.). Bilbao: Publicaciones de la Universidad de Deusto. ゴンザレス, J.・ワーヘナール, R. (2012) 『欧州教育制度のチューニング—ボローニャ・プロセスへの大学の貢献』(深堀聰子・竹中亨訳) 明石書店.
- Jankowski, N. A., & Marshall, D. W. (2017). *Degrees that matter: Moving higher education to a learning systems paradigm*. Sterling, VA: Stylus.
- Kosslyn, S. M., & Nelson, B. (Eds.). (2017). *Building the intentional university: Minerva and the future of higher education*. Cambridge, MA: The MIT Press.
- 松下佳代（2019a）「学習成果とその可視化」中央教育審議会大学分科会教学マネジメント特別委員会（第6回）. (https://www.mext.go.jp/content/1417855_002.pdf)（2019年12月22日アクセス）
- 松下佳代（2019b）「汎用的能力を再考する—汎用性の4つのタイプとミネルヴァ・モデル—」『京都大学高等教育研究』25，67-90.
- Matsushita, K., Ono, K., & Saito, Y. (2018). Combining course- and program-level outcomes assessments through embedded performance assessments at key courses: A proposal based on the experience from a Japanese dental education program. *Tuning Journal for Higher Education*, 6 (1), 111-142. (doi.org/10.18543/tjhe-6 (1)-2018 pp111-142)（2019年12月22日アクセス）
- 小田陽平・小野和宏・藤井規孝・小林正治・前田健康（2017）「診療参加型歯科臨床実習におけるweb公開型電子ポートフォリオの開発と運用」『日本歯科医学教育学会雑誌』33，65-73.
- 小野和宏・松下佳代（2015）「教室と現場をつなぐPBL—学習としての評価を中心に—」松下佳代・京都大学高等教育研究開発推進センター編著『ディープ・アクティブラーニング—大学授業を深化させるために—』勁草書房，pp.215-240.
- 小野和宏・松下佳代（2016）「初年次教育におけるレポート評価」松下佳代・石井英真編『アクティブラーニングの評価』東信堂，pp.26-43.
- 小野和宏・松下佳代・斎藤有吾（2014）「PBLにおける問題解決能力の直接評価—改良版トリプルジャンプの試み—」『大学教育学会誌』36(1)，123-132.
- 小野和宏・斎藤有吾・松下佳代（2018）「PBLを評価する改良版トリプルジャンプにおける『学習としての評価』の要因」『京都大学高等教育研究』24，35-43.
- 佐藤学（1996）『教育方法学』岩波書店.
- Suskie, R. (2009). *Assessing student learning: A common sense guide* (2nd ed.). San Francisco, CA: Jossey-Bass. サスキー, L. (2015). 『学生の学びを測る—アセスメント・ガイドブック—』(齋藤聖子訳) 玉川大学出版部.
- WSCUC. (2015). *Using evidence in the WSCUC accreditation process: A guide for institutions* (2nd ed.). (<https://wascsenior.box.com/shared/static/5fum7qzstgwx0h74kae2b9zbss22gyaj.pdf>)（2019年11月12日アクセス）