

# 学生に求められる能力とその評価

松下 佳代

京都大学 高等教育研究開発推進センター教授

本日は、「学生に求められる能力とその評価」についてお話します。まず、3つの問いを立てました。「学生に求められる能力とはどういうことか」、「その能力はどう形成されるのか」、「その能力はどのように評価されるのか」の3つです。本日は、1番目と3番目を中心にお話します。

## 1. 能力への注目

1998年にベストセラーになった『老人力』は、否定的に見られやすい「老人」の持つ力を、ポジティブに評価するという非常に面白い本でした。その後も、2007年の『鈍感力』、2008年の『悩む力』、2012年の『聞く力』と続けて「何々力」や「何々する力」というのが、ベストセラーになりました。これ以外にも、「何々力」「何々する力」「何々能力」などが、巷に溢れています。このような現象が、90年代の半ば辺りから見られるようになってきたことが、能力が注目される一つのきっかけだったと思います。

教育の分野でも、90年代以降、日本では2000年代に入ってから特にそうですが、さまざまな形で「何々力」「何々能力」が提唱されるようになってきました。私は、それらを総称して〈新しい能力〉と名付けました。〈新しい能力〉とは、ポスト近代社会や後期近代社会と言

われる社会において求められるようになった能力のことです。

では、なぜ90年代以降に、こういう〈新しい能力〉が氾濫するようになったのかと言うと、ポスト近代社会や後期近代社会になって、これまでの生き方の定番が揺らいできたからではないでしょうか。これまでは、「定番」に頼って生きていればよかったところが、人々が自らの力で人生のさまざまな局面を切り開いていかなければならなくなった時代、だからこそ教育の世界でも教育以外の世界でも、「何々力」「何々能力」が求められるようになってきたのではないかと思います。

図1は、日本の90年代半ば辺りからの「何々力」「何々能力」を一覧にしたものです。2014年12月に出された高大接続答申では、「生き

日本の場合

【初等中等教育】		
生きる力	文部科学省	1996年
リテラシー	OECD-PISA	2001年～
学力の三要素	学校教育法	2007年
21世紀型能力	国立教育政策研究所	2013年
【高等教育】		
就職基礎能力	厚生労働省	2004年
社会人基礎力	経済産業省	2006年
学士力	文部科学省(学士課程答申)	2008年
ジェネリックスキル	OECD-AHELG	2012年
【労働政策・成人一般】		
エンプロイアビリティ	日本経営者団体連盟(日経連)	1999年
成人力	OECD-PIAAC	2012年
【全般】		
キー・コンピテンシー	OECD-DeSeCo	2003年
人間力	内閣府(経済財政諮問会議)	2003年
基礎的・汎用的能力	文部科学省(キャリア教育答申)	2011年

図1

る力」や「学力の三要素」が高大接続、つまり、中等教育と高等教育の両方に一貫する必要な能力として取り上げられています。したがって、この表の中では初等中等教育と高等教育を区別して書いていますが、そこに繋がりが出てきているのが最近の動向だと言えるでしょう。

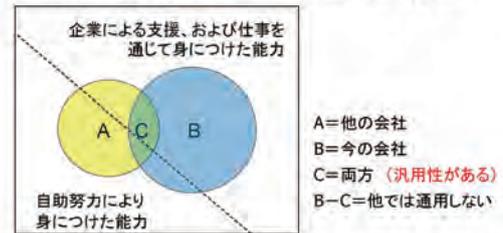
高等教育では、若者の就職難という問題に対して、厚生労働省がもっとも早く「就職基礎能力」の必要性を訴えました。しかし事業仕分けで事業が廃止されたため、現在の大学教育では、経済産業省の「社会人基礎力」や文科省の「学士力」の方が注目されています。世界的なところでは、OECD-AHELOの「ジェネリックスキル」などがあります。また、労働政策の中では「エンプロイヤビリティ」という概念がかなり早い段階で提唱されましたが、これは高等教育の中でも使われる概念になっています。さらに、年齢を限定せずに「キー・コンピテンシー」、「人間力」、「基礎的・汎用的能力」などが提唱されています。

先ほど、「定番が揺らいできたことが『何々力』『何々能力』という言葉の氾濫につながっている」と申し上げましたが、その一つの例として「終身雇用制の揺らぎとエンプロイヤビリティ」を挙げてみます。99年に日経連がエンプロイヤビリティという言葉を使いましたが、その背景にあったのは終身雇用制の揺らぎでした。

では、エンプロイヤビリティがどういう概念かについて説明します。図2は、日経連のエンプロイヤビリティのモデルです。これまでの終身雇用であれば、このBの部分だけで十分でした。このB、今いる会社の中で必要な能力、たとえ配置転換などがあっても、その会社で必要とされている能力を身につけていればよかったということです。ところが、労働移動が盛んになると、このCの部分がある人はいいいのですが、Cの部分がない、即ちこの会社でしか通用しない能力しかないのは困るので、たとえ

## 定番の揺らぎ と 能力

### ● 終身雇用制の揺らぎ と エンプロイヤビリティ



### ● エンプロイヤビリティ

→ 就職基礎能力、社会人基礎力、汎用的能力

図2

今の会社から移っても通用するような能力を身につけていくために、Aの部分が必要となってきます。しかも、その能力はこれまでのように企業がお金を出した研修やOJT（オンザジョブトレーニング）の中で身につけさせていくのではなく、自助努力で身につけていきなさいというモデルなのです。そして、Cの部分は、さまざまな会社で必要となるため汎用性があるということになります。これが、「汎用的能力」や「ジェネリックスキル」と言われている背景の一つになっているものだと思います。

実際、「就職基礎能力」にあたる英語として、厚生労働省は“employability”を用いていました。社会人基礎力やキャリア教育答申の汎用的能力も、このようなエンプロイヤビリティの考え方を基盤にしていると思います。

このように、社会の中の定番の揺らぎが能力の氾濫につながっているわけです。

では、このような〈新しい能力〉にはどんな特徴があるのでしょうか。ここでは3点挙げておきます(図3)。

まず、特徴的なのは、「対象・範囲の広さ」です。ここまで説明してきたのは日本の場合ですが、似たような概念が世界中で見られます。それは、日本が世界の動向と並行して教育改革を進めてきているということでもあるわけです。

## 〈新しい能力〉の特徴

- 対象・範囲の広さ
  - 年齢…小学生から社会人まで
  - 国・地域…世界中で
- 能力の中身の広さ
  - 認知的側面(知識や技能など)だけでなく、**情意的側面**(態度など)や**社会的側面**(対人関係など)も含まれる
  - to know と to do  
(単に、知っているだけではなく、それを必要な場面で使えること)
- 教育への影響の大きさ
  - 教育目標であるだけでなく、**評価対象**として設定されている

図3

次に、「能力の中身の広さ」です。いわゆる、知識・技能など認知的な面だけではなく、態度・意欲といったような情意的な側面や、あるいは社会的な側面として対人関係・コミュニケーション、このようなものも含めて能力が捉えられるようになってきました。それは単に知っているだけではなく、それを必要な場面で使えることが求められるようになったことを表しています。実際に、社会に出て何かをしようとすると、一人ではなく、何人かの人達と共同でコミュニケーションを取りながら進めていくこととなります。しかもそれには、ある一定の意欲、それに向かう方向性・ベクトルが必要になってきます。そのようなところから、単に認知的な側面だけではなく、情意的側面や社会的側面も求められるようになってきたのだと言えるでしょう。

それから、「教育への影響の大きさ」として見逃せないのが、能力が教育目標であるだけでなく、評価対象としても設定されていることです。創価大学のAPでも「学修成果の可視化」が謳われていますが、可視化をして評価をするわけで、評価されるということは学生にも教員にも大きな影響を与えます。これも世界中で見られる特徴です。

## 2. 学生に求められる能力

今まで前提的なことをお話ししてきましたが、「学生に求められる能力」とはどのようなものかについて、特に影響力のある能力概念を5つほど挙げます(図4)。

### とくに影響力のある能力概念

- 日本
  - 経産省: 社会人基礎力
  - 文科省: 学士力
  - 日本学術会議(分野別参照基準):  
分野に固有の能力とジェネリックスキル
- 海外
  - AAC&U(全米大学・カレッジ協会): 本質的学習成果
  - Tuning Project: 一般的コンピテンスと分野別コンピテンス

図4

例として、経産省の「社会人基礎力」文科省の「学士力」を挙げますと、分野別質保証のため、教育課程編成のための参照基準が日本学術会議の分野別分科会の中で進められています。そこでは、分野に固有の能力とジェネリックスキルを分野ごとに設定しなさいということが言われています。

海外でも多くの例がありますが、特に日本の教育に対して影響力の大きいものとしては、AAC&U(全米大学・カレッジ協会)の「本質的学習成果(Essential Learning Outcomes)」という考え方があります。それから「Tuning Project」はヨーロッパから始まり、今では世界中に広がっているものですが、この中では、一般的なコンピテンスと分野別コンピテンスという概念が抽出されています。これらはお互いに関係があって、例えばAAC&Uの本質的学習成果と学士力は非常に近い関係にあります。日本学術会議の考え方は、Tuning Projectの考え方と非常に似ています。

「社会人基礎力」について、「アクション」「シンキング」「チームワーク」の3つはよく取り上げられますが、そこには知識が全く入っ

ていないように見えます(上のモデル図)。ただし、もう一つ大きな枠組(下のモデル図)で見ると、「社会人基礎力」というものは「基礎学力」や「専門知識」と相互作用しながら培われていくものであり、「基礎学力・専門知識を活かす力」ということになっています。この能力と知識の関係を見落とさないようにしないとイケません(図5)。

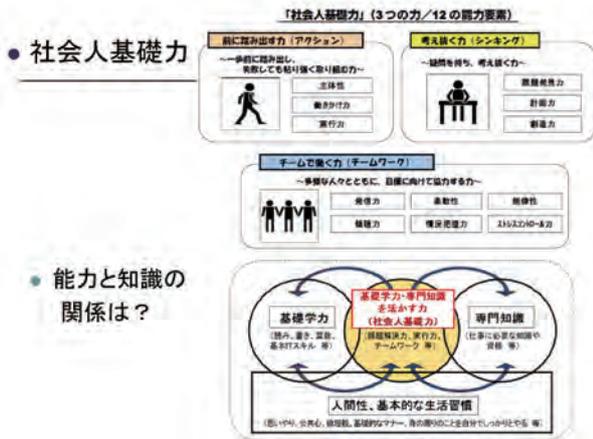


図5

注：図は、経済産業省「社会人基礎力」のウェブページ (<http://www.meti.go.jp/policy/kisoryoku/>) より引用。

図6は、本質的学習成果と学士力を比較した表ですが、両者を並べてみると、4つの大きなカテゴリーも、その中のサブカテゴリーも非常に似通っていることがわかります。学士力を見たときに「統合的な学習経験と創造的思考力」というものがつかみにくかったのですが、本質的学習成果の方では、「統合的学習」となっ

● 本質的学習成果と学士力

本質的学習成果 (AAC&U)	学士力 (文科省)
<b>人類の文化や自然界についての知識</b> ・科学、数学、社会科学、人文学、歴史、言語、芸術などの学習を通じて	<b>知識・理解</b> ・多文化・異文化に関する知識の理解 ・人類の文化、社会と自然に関する知識の理解
<b>知的・実践的スキル</b> ・探究と分析 ・批判的思考・創造的思考 ・文章コミュニケーション・口頭コミュニケーション ・量的リテラシー ・情報リテラシー ・チームワークと問題解決	<b>汎用的技能</b> ・コミュニケーションスキル ・敬重のスキル ・情報リテラシー ・論理的思考力 ・問題解決力
<b>個人的・社会的責任</b> ・市民としての知識と関与(ローカル、グローバルに) ・異文化の知識と能力 ・倫理的な推論と行為 ・生涯学習のための基礎とスキル	<b>態度・志向性</b> ・自己管理力 ・チームワーク ・リーダーシップ ・倫理観 ・市民としての社会的責任 ・生涯学習力
<b>統合的学習</b> ・一般教育・専門教育での統合とより高度な達成	<b>統合的な学習経験と創造的思考力</b>

(AAC&U, 2007; 中教審, 2008)

図6

いますね。AAC&Uでは、例えばキャップストーン科目という、日本でいう卒業研究のような科目で、この統合的学習を行うことになっています。そう考えると、「統合的な学習経験と創造的思考力」の意味するところもわかりやすくなります。いずれにせよ、文科省の学士力が、AAC&Uの本質的学習成果という考え方に強く影響を受けて提案されていることがわかりいただけるかと思います。

では、なぜ学習成果が目標になるのでしょうか。これも2000年頃からの傾向だと思いますが、ここでいう学習成果は実際に達成された学習成果ではなく「期待される」、英語でいう expected や desired という形容詞が付く学習成果のことです。「期待される」、「求められる」学習成果。つまり、実際に学習した結果というよりは、むしろ学習の成果を生み出すようなものとして、目標に掲げられています。

日本学術会議の分野別質保証のための参照基準(図7)には、1番目、2番目が身に付けるべき基本的な素養ということで、知識・理解と能力が挙げられ、さらに能力には、分野に固有の能力とジェネリックスキルが入っています。3番目の学習方法と評価方法についても、それぞれの分野で基本的な考え方を決めるよう促しています。4番目は、専門教育と教養教育との関わりです。このように、知識以外と能力、能力も分野別の能力だけではなく、ジェネリックスキルも含めて各分野で設定するようにとしてい

● 日本学術会議：分野別質保証のための参照基準

1. 当該学問分野の定義と固有の特性
2. 当該学問分野で学生が身に付けるべき基本的な素養  
 (1) 基本的な知識と理解  
 (2) 基本的な能力: 分野に固有の能力 と ジェネリックスキル
3. 学習方法と学習成果の評価方法に関する基本的な考え方
4. 市民性の涵養をめぐる専門教育と教養教育との関わり

(日本学術会議「回答 大学教育の分野別質保証の在り方について」2010.7)

図7

るわけです。

これと非常に近い関係にあるのが、Tuning Projectです。チューニングとは、オーケストラのチューニングのように、国境を越えた大学間で、さまざまな楽器が一つの曲を奏できるようにお互いに調整しながらカリキュラムの枠組みを共有していこうという考え方です。元々はヨーロッパで始まり、今では世界中に広がっています。2010年当時で53カ国でしたが、現在は日本でも、Tuning Japanとして国研にチューニング情報拠点が立ち上がっています。

Tuningでは、目標を能力(コンピテンス)として抽出し、それに基づいてカリキュラムを作り、授業学習を行い、そして評価をします。このようなサイクルで質保証を行っていくという流れが作られています。ここでの能力(コンピテンス)とは、各大学の学位プログラムを履修した総合的な学習成果として、学生が獲得することが期待されている、知識・技能・態度が有機的に結合したものを指します。

コンピテンスは、知識と並列する概念として使われることもあります。ここでは、知識を含み、さまざまな要素が有機的に結合したものと捉えられています。では、どのようにコンピテンスを抽出していったのかというと、大学教員、卒業生、雇用主の三者に質問紙調査を行い、上位にリストされたものを選ぶというやり方で、三者異なる利害を持った人達の意見を集約する

Label	Description	Academics	Graduates	Employers	Grad & empl.
imp5	Grounding in basic knowledge of the profession	8	11	14	13
imp6	Oral and written communication in your native language	9	7	7	5
imp7	Knowledge of a second language	15	14	15	15
imp8	Elementary computing skills	16	4	10	8
imp9	Research skills	11	15	17	16
imp10	Capacity to learn	3	2	1	2
imp12	Critical and self-critical abilities	6	10	9	10
imp13	Capacity to adapt to new situations	7	5	4	4
imp14	Capacity for generating new ideas (creativity)	4	9	6	7
imp16	Decision-making	12	8	8	9
imp18	Interpersonal skills	14	6	5	6
imp28	Ethical commitment	13	16	13	14

・一般的コンピテンスは、大学教員、卒業生、雇用主への調査をもとに抽出

Gonzalez & Wagenaar (2008, p.47)

図 8

形で、一般的なコンピテンスが抽出されています。

ここまで、欧米と日本の高等教育で、〈新しい能力〉についてどんな議論が行われているかを見てきましたが、ここで「学生に身につけさせたい能力」を考える際の3つの論点を挙げたいと思います。1番目は、「知っているだけでなく行えること」ことです。これに関しては、カナダの高等教育の研究者であるジェームズ・コテ氏が「アイデンティティ資本モデル」(図9)を挙げています。これまでは人的資本と社会関係資本が重要とされてきましたが、「エイジェンティック・スキル」という行為主体的な能力も必要になってきており、この後期近代社会を生きていくためには、このような能力も必要になっているという議論が展開されています。

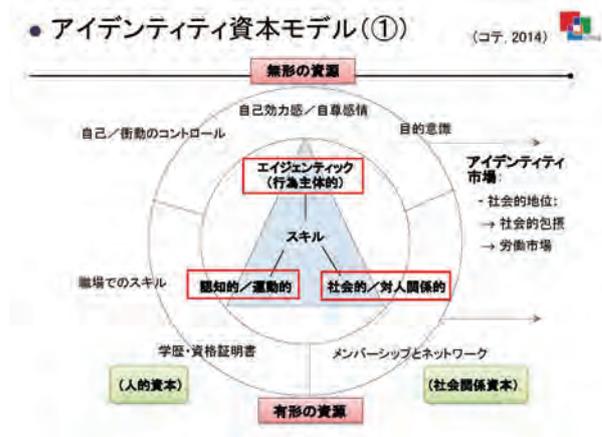


図 9

2番目は、一般的能力と分野固有の能力の関係です。Generic(一般的)というとなんかぼんやりとしていますが、GenericはむしろTransferable(転移可能)と考えたほうが良いと思います。具体的にはこのようなモデルで考えています。問題解決能力を例にとると、各分野で特徴のある問題解決能力というものがあります。その分野固有の特徴と、ある程度分野を超えた共通性があるわけです(図10)。学生たちは、分野の中での問題解決能力を身につけな

がら、その分野独自の特徴と分野を超えた共通性の部分の両方を意識・認識することによって、異なる分野・文脈にも転移させていけるようになります。このようにして、徐々に自分の能力の適用範囲を広げていくと考えると、Genericが曖昧な能力の概念ではなく、実際に学ぶことと結びつけて考えやすくなると思います。

### ● 一般的能力と分野固有の能力の関係(②)

- 同じ問題解決能力でも、  
 <分野固有の特徴>と<分野を超えた共通性>がある
- その両方を認識することで、転移させやすくなる

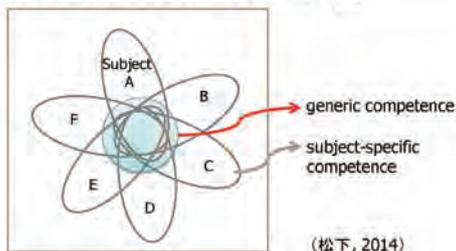


図 10

3番目は、汎用性をもつ深い知識を身につけるといことです。〈新しい能力〉論の多くでは、「ジェネリックスキル」に見られるようにスキルの汎用性に目が向けられることが多いのですが、私は、「スキル」だけでなく「知識」についても、他の場面に転移できないものと、非常に広がりのある、さまざまな場面で使えるものがあると考えています。小田中直樹氏は著作の中で、「生きる力」という言葉を使いながら、さまざまな社会の現象を捉えるための、まさに「生きる力としての経済学」の概念があると述べています。例えば「分配」や、今であれば「格差」などは、さまざまな現象に幅広く使える、転移可能な概念だと言うことができるのだと思います。

McTighe & Wiggins (2004) は、このような知識とスキルに見られる深さの軸を「知の構造」という図式で表しています。例えばアメリカ史で第二次大戦を教えるときに、「ヒトラーがドイツで政権を取ったのは何年だった」のような事実に基づく知識があります。そして、より深い

ところへ掘り下げていくと、「戦争の中には“正義の”戦争とみなされるものがある」というような一般化された命題があります。このロジックは、ベトナム戦争の時にもイラク侵攻の時にも使われてきましたし、現在であればテロとの戦いでも使われています。このように、一般化された原理や概念を深く理解した学生であれば、今後似たようなことが起きたときに、「これにはあれが使えるのではないか」と自分の見方を広げていけます。そのような力を、大学生であればぜひ身につけてほしいと思います。

このような知識は、理解するのは大変ですが、いったん理解してしまえばなかなか忘れないものです。私は「忘れ残り」と呼んでいます。多くの細かいことは忘れてしまっても、根本の考えの部分は忘れないで残っているものです。学生には、そのような知識を是非授業の中で身につけてほしいと思います。

### 3. 学習評価の構図

さて、ここまでは、現在、世界的に見られる〈新しい能力〉の流れをふまえながら、学生に求められる能力についてお話ししてきました。ここからは、では、そうした能力をどう評価するのか、についてお話ししていきたいと思ます。

学習成果への注目が特になされるようになってきたのは、2008年の学士課程答申あたりからだと思ます。学習成果に注目するということは、「教員が何を教えた(つもり)か」から、実際に「学生が何を学んだか」に目を向けるということです。そうすると、学習成果をどう見るか、どう評価するかが問題になってきます。現在行われている学習評価は、「直接評価」と

「間接評価」、それから「心理測定学的パラダイム」と「オルターナティブ・アセスメントのパラダイム」、この2つの軸で捉えることができます。2012年の中教審の「質的転換答申」で、アセスメント・ポリシーという言葉が使われていて、さまざまな測定の手法が例示

されていますが、それもこの2つの軸からできる4象限の中に位置づけることができます(図11)。学修行動調査は第II象限、アセスメントテストは第III象限、ルーブリックや学修ポートフォリオは第IV象限に当たります。

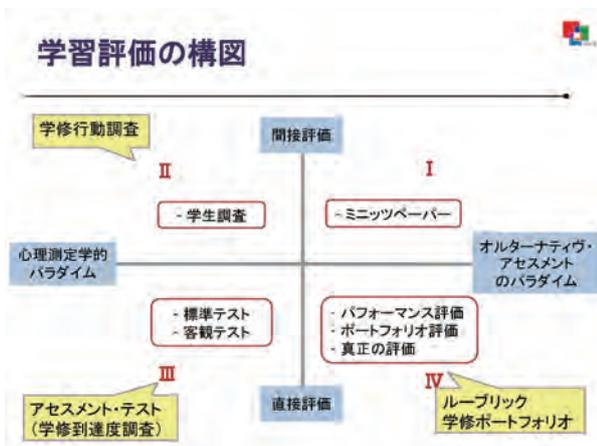


図 11

各象限の具体例をタイプ別に挙げると、ミニツペーパーは、タイプIです。学生調査、例えばさまざまな能力についてのどの程度身についているかを5件法などで答えさせる調査などは、タイプIIです。センター試験は、タイプIIIです。レポート・卒業論文・プレゼン・口頭試問・プロジェクト・卒業制作・模擬授業・模擬裁判などはタイプIVにあたりますが、大学の中では学習成果を測る重要なエビデンスになってきます。これらのものをどのように評価すればよいのか、どうデザインすればよいのかが、この後のメインテーマになります。

まず、直接評価－間接評価という軸から見いきましょう。両者の違いを一言で言うと、「間接評価」は、学生が自分は何ができてと思っているか、さまざまな能力について自分はどのぐらいできていると思うかを彼らに答えさせるのが特徴です。一方、「直接評価」は、実際に何か知識・技能を表出させて学生の学習成果を直接に評価する、実際に何ができて見えるものです。何ができてと思っているのかという学生の自己認識を通して評価をすることが

間接評価であり、表出させたものを評価するのが直接評価です。

もう一つの軸は、「評価の2つのパラダイム」(図12)です。

	心理測定的パラダイム	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、知能理論	構成主義、状況論
評価目的	アカウンタビリティ、質保証 プログラム評価	指導、学生の成長 授業評価
評価項目	分割可能性	複合性
評価機能	総括的評価	形成的評価
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	客観テスト(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

図 12

特に、評価目的・評価基準・評価データの違いを見ると、両者の違いがよくわかります。例えば、標準テストや学生調査のように定量的なデータが出てきて、客観的に評価ができるものは、集合体で利用されることが多く、アカウンタビリティのために使われたり、あるいはプログラム評価のために使われたりします。ただ、それでは一人ひとりの学生の成長を見たり、指導したりすることは難しくなります。そこでは、質的データも用いながら、個々の実践者(教員や学生自身)が、主観的ではあっても恣意的にならないよう間主観性をもった評価を行うことが必要になってきます。

#### 4. パフォーマンス評価とルーブリック

間接評価・直接評価と2つのパラダイムは、掛け合わせるとさまざまなタイプの評価の違いが見えてきます。

なかでも、本日焦点を当てるのは、評価方法のうち、真正の評価やポートフォリオ評価、パフォーマンス評価と言われるものです。真正の評価やパフォーマンス評価について一番わかりやすいのは、自動車の運転免許が挙げられます。自動車の運転免許、例えば仮免でも実際に運転

させてみて、その運転能力を見ます。このように、実際に何かを遂行させたり、卒業制作であれば実際にプロダクトを作らせたりして、それを評価する方法、つまり、パフォーマンス(実演や作品)を通じて評価する方法が「パフォーマンス評価」です。その中でも、例えば仮免と本試験を比べると、本試験は路上検定で行われ、より本物らしさが高まっています。パフォーマンス評価の中でも、シミュレーション的に行う場合もあれば、より本物に近い場面で行う場合もあります。パフォーマンス評価の中でもより本物らしさを備えたものを「真正の評価」といいます。実は、論者によって概念の説明の仕方や関係付けの仕方が異なりますが、私は、この説明の仕方が一番わかりやすいと思っています。また、学習のエビデンスになるものは他にも客観テストやワークシートなどさまざまあります。それらを全部束ねてポートフォリオにして、どのように自分が学習者として成長してきたのかを振り返りながら評価する、それが「ポートフォリオ評価」です。

なぜか日本の大学教育の中では「ルーブリック評価」という言葉が頻繁に使われています。しかし、これは日本の大学教育に特異な現象で、アメリカでは「ルーブリック評価」という言葉はほとんど使われません。なぜかと言うと、「ルーブリック」というのはオルターナティブ・アセスメントの評価基準にすぎないからです。私は、ルーブリックだけ一人歩きさせるのではなく、あくまでもパフォーマンス評価の中で、そのルーブリックをどう使うかということ発想すべきではないかと考えています。では、パフォーマンス評価とは何か、どのように作るのかということになりますが、これをモデル化したものが、図13になります。

能力については、それ自体を見ることはできません。例えば、力持ちという場合でも、普通の人を持ち上げられないような重たい物を持ち上げられるという観察可能な事象にして、それを解釈する形で、「この人はすごい力持ちだ」

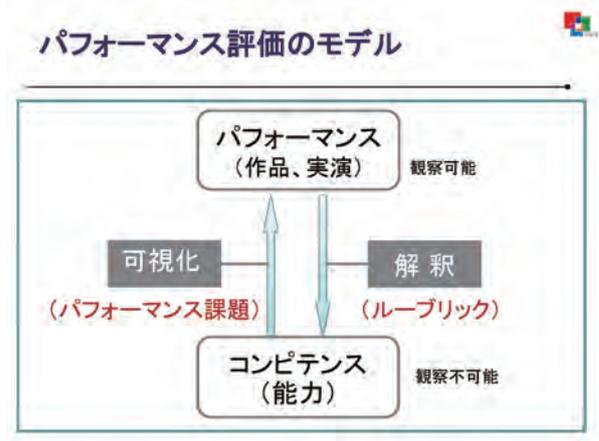


図 13

と言っているわけです。したがって、能力を評価するという事は、それ自体、その能力そのままを評価することはできないので、何らかの事柄を行わせて、その解釈をして、能力があるかないかを判断しています。パフォーマンス評価において、この能力を可視化するところで使われるのが「パフォーマンス課題」であり、その可視化されたパフォーマンスから、そこにどんな能力が表れているのかを見るときに使われるのが、「ルーブリック」ということになります。

大学教育はパフォーマンス課題の宝庫だと思います。ですが、評価基準については、ほとんど教員の主観に委ねられてきました。ルーブリックというのは、パフォーマンスの質を段階的・多面的に評価するための評価基準表ですが、これは、大学教員のようなエキスパートが持っている鑑識眼、見る眼を明示化して、複数の教員間や、教員と学生の間、あるいは教員と他の部外者との間で共有できるようにするためのツールです。普通これは、図14のようなマトリックス形式で書かれることが多いです。

元々は質的なデータでも、このルーブリックを使うことで、量的にも表現できる可能性が生まれてくるということで、ルーブリックはかなり使われるようになってきているのだと思います。質的転換答申でも言及されましたので、それによって拡がってきていますが、元々は政策云々ではなく、高次の能力を学生に身に付けさ

## ルーブリック

### ● ルーブリックとは

- =パフォーマンスの質を段階的・多面的に評価するための評価基準表
- 「専門家の鑑識眼」を明示化し、共有できるようにするツール
- パフォーマンスの質を量的表現に変換する働きも

### ● ルーブリックへの注目

- 高次の能力を評価 ……正誤ではない、出来ぐあいの連続体

The diagram shows a rubric table with 4 levels (レベル4 to レベル1) and 4 criteria (観点1 to 観点4). The criteria are listed on the left, and the levels are listed at the top. A red box highlights the '記述語 (descriptor)' column, and another red box highlights the '観点1 (規準)' row.

図 14

せなければならない、その能力をどう評価するかといったときに、あまりできないところからよりよくできる、見事にできるというような、能力や出来具合の幅、連続体の中で、ある段階を区切ってその質的差異をとらえようとするもの、それがこのルーブリックです。

ルーブリックの例はさまざまありますが、ここでは、AAC&UのVALUEルーブリック、関西国際大学のコモングルブリック、そして山口大学、新潟大学で開発されたルーブリックをご紹介します。VALUEルーブリックは、大学間で科目横断的に共有されており、関西大学のコモングルブリックは学内で科目横断的に共有されています。また、私も開発に関わった山口大学や新潟大学歯学部の場合は、特定の科目において教員間で共有されています。このようにそれぞれのルーブリックがどう共有されているかという点で見ると、それぞれ階層が異なります。

まず、VALUEルーブリックは、本日最初の方でお話しした「本質的学習成果」を評価するためのルーブリックです。VALUEというのは、Valid Assessment of Learning in Undergraduate Education (学士課程教育における妥当な学習評価) のことですが、これは何に対して言っているのかといえば、単にリテンション率や標準テストではない、本当に妥当な評価という意味をこめて、validと言っているのです。

VALUEルーブリックには、批判的思考のルーブリックや倫理的推論のルーブリックなど16のルーブリックがあります。これは特定の科目のために使われるルーブリックではなく、4年間かけて使われる長期的なルーブリックです。また、VALUEルーブリックは、1300校以上あるAAC&Uの加盟校に共通のルーブリックとして提案されています。そのため、実際に使うときには、自分の大学・部局や科目に合わせてローカライズすることが必要になりますが、そのプロトタイプとなるような、大学間で科目横断的に共有されるものとして作られたのがこのVALUEルーブリックです。

例えば図15は、文章コミュニケーションVALUEルーブリックですが、大まかに言って、「ベンチマーク」というのは1年生、「キャップストーン」は4年生を指しています。

【例】文章コミュニケーション VALUEルーブリック

	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
文章作成の 文脈と目的	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	適切に関連性があり読者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、作品全体を形づけている。	.....	.....	適切に関連性のある内容を用いることによって、作品の何が所か、シンプルなアイデアを展開している。
ジャンルと 学問分野の 約束事	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、細かい注意を向けようとしている。	.....	.....	基本的構成や提示のしかたについて一貫した体系を築こうとしている。
資料と根拠	当該の学問分野やジャンルにふさわしいアイデ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。	.....	.....	アイデアを築くために、資料を扱おうとしている。
構文と技法を 操ること	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。	.....	.....	用語法に誤りがあるために、意味の伝達が行われないような言葉遣いをしていいる。

(注) レベルを満たさないものはレベル0とされるので、実質的には5段階のレベルからなる。

図 15

1年生と4年生での変化を見ることで、学生がどれだけ成長したか、逆に言うと大学はどれだけ学生を成長させることができたかを見られるように作られています。

図16は、バイオロジーの1年生の科目で使われたルーブリックです。いくつかのVALUEルーブリックから自分の科目に合わせて段階設定をして、得点化して使われています。このように自分の大学・部局・科目に合わせてアレンジして使うわけです。

関西国際大学のコモングルブリックは、ライ

【例】VALUEルーブリックを利用した  
レポート評価用ルーブリック (BIO109)

● VALUEルーブリック → 大学・学科・科目にあわせて修正

	2	1	0	得点
文章作成の目的 (文章コミュニケーション)	読者を考慮していること、および、研究の焦点が明確であることが、タイトルに示されていた。	.....	.....	
問題の特定 (文脈的能力)	研究の主要問題が一般的な読者に関心のあるものになっていた。 .....	.....	.....	
エビデンス (批判的思考)	ソース(情報源)からの情報を評価し、自分自身の解釈を生み出していた。 .....	.....	.....	
問題の説明 (批判的思考)	すべての問題が記述され、明確化されていた。 .....	.....	.....	
構文と技法を操ること (文章コミュニケーション)	使用された言葉は、読者に意味を明確に伝えている。 .....	.....	.....	(Rhodes, 2013)

図 16

ティング、プレゼンテーション、リサーチという、さまざまな科目で共通して必要になるであろうスキルを取り上げ、それに6段階に設定し、下位学年用と上位学年用に分けて使われています。段階が0から6に分けられ、0から3までが下位学年用、3から6が上位学年用になっています。全体を通して、1年生から4年生を終える時までの成長が見られるようになっています。

山口大学のコモンルーブリックは、全学必修の初年次科目「山口と世界」で使われています。統一科目名で、統一の目標を立てていますが、実際にはそれぞれの教員が自分の専門を活かしながら担当するので、学生に調査をさせる人もいれば、研究に近いことをさせる人もいれば、商品開発をさせる人もいます。そういう多様性を包含できるような形で作られているのがコモンルーブリックです。それぞれの教員が自分の担当する「山口と世界」の特徴に合わせて修正できるようになっています。

「山口と世界」コモンルーブリックの規準は、「発見する、はぐくむ、かたちにする、分かちあう、振り返る」という学習活動のプロセスに沿って作られています(図17)。

「分かちあう」では、例えば商品開発であれば、実際開発した商品を地元の人たちに味わってもらったりしながら、そこから評価を受けます。単に教員や学生仲間からだけでなく、部外者からも評価を受けるように、評価を外に開くことも行われているわけです。

● 「山口と世界」コモンルーブリック

規準	内容	レベル3 【秀】	レベル2 【良～優】	レベル1 【可】	レベル0 【不可】	評価 比率
発見する	山口に関連するテーマ設定、企画立案	山口に関連する課題に応じて、十分な下調べもとづき、適切かつ独創的なテーマ設定、企画立案ができる	山口に関連する課題に応じて、下調べをした上で、テーマ設定ができる	山口に関連する課題に応じたテーマ設定ができる	レベル1に満たない	20%
はぐくむ	テーマ設定、企画にもとづく情報収集およびコミュニケーション					20%
かたちに する	構案、作品化、発表資料、レポート等、プロダクツの作成					30%
分かち あう	公開、プレゼンテーション、チームワーク					15%
振り返る	他者および自分(たち)の企画およびプロダクツの評価					15%

図 17

5. パフォーマンス評価の開発事例

ここまで、VALUEルーブリック、関西国際大学のコモンルーブリック、山口大学のルーブリックについてご紹介してきましたが、新潟大学歯学部のコモンルーブリックを含めたパフォーマンス評価の開発には、私もかなり深く関わったので、もう少し詳しくお話ししていきましょう。新潟大学歯学部では、目標として「新潟大学歯学部版学士力」を設定し、その評価については、レポート評価とPBLテュートリアルのための評価を開発しています。「新潟大学歯学部版学士力」では専門的能力と汎用的能力、知識・理解、態度・姿勢の全部で25項目挙げられていますが、このような能力を育てる上で重要な取り組みになっているのがPBLテュートリアルです。

PBLテュートリアルの進め方ですが、まず、ある事例をシナリオで読ませます。最初の段階では学生はまだ知識が限られているので、一応解決策を立てますが、さまざまわからないことが出てきます。そこで学習課題を設定します。ここまですべて最初の1時間で行います。その次の週の授業までに、それぞれ個別で学習課題について調査をします。1週間後に、それぞれが持ち寄った知識を統合して、最初の仮説でよかったのかを検証して、自分たちの解決策を立て直していきます。このプロセスには教員がテューターとしてつきます。また、個人学習とグループ学習、教室内と教室外の学習を組み合わせ

いることも特徴です。

問題解決には知識も必要なので、新潟大学では1週間の授業の中に、PBLテュートリアルとそれに関係する講義や演習など、授業形態の異なるものがうまく組み合わされています。そのため、あるPBLを行う時には、それに必要な知識に関する内容を講義で扱うようになっていきます。

用いられるシナリオは例えばこんな感じです。あなたは新潟大学医歯学総合病院の歯科衛生士で、タバコが生きがいが、歯周病を発症して糖尿病もあるという患者さんに禁煙指導をすることになったというような設定です。糖尿病があると、歯周病は悪化しやすいらしいですね。そういうことも学生たちはまだ知らないので、学習課題を立てて、自分で調べながら、どうやってタバコが生きがいの人に禁煙指導をしていくかを考えていきます。

では、PBLテュートリアルがめざす問題発見・解決能力や自己学習能力、知識の統合的理解などをどのように評価したかについてですが、大きく3段階に分けて評価しました(図18)。もともとPBLではトリプルジャンプという方法がありましたが、これはその改良版で「改良版トリプルジャンプ」と名づけています。問題発見から解決策の提案までにあたるステップⅠ、Ⅱは、ワークシートを使って筆記課題で評価します。解決策の実行にあたるステップⅢは、解決策をロールプレイで実演してもらって、その

場でフィードバックもします。

このようなプロセスを評価するために、ステップⅠ・Ⅱ用まとめて一つ、ステップⅢ用でもう一つ、ループリックを作成しました。きょう前半の方で、問題解決能力を例に挙げて、「同じ問題解決能力でも、＜分野固有の特徴＞と＜分野を超えた共通性＞がある」というお話をしました。例えば、このPBLで開発したループリックについて、どういうところに学問分野の独自性が入っているかということ、やはり対人相手の仕事ですので、問題解決をするというのは他者に働きかける形でなされます。共感的態度やコミュニケーション、あるいは患者さんとのやりとりの中で情報を絶えず取り入れ、情報を統合しながら解決していかなければいけないというところに、その分野独自の特徴が表れています。しかし逆に言うと、それ以外のところは、問題発見から解決策の実行、結果の評価にいたるまで、かなり問題解決一般に通ずる特徴もあります。このようにして問題解決能力をとらえて、ループリックを作成しました。

図19は、学生の変化を表したグラフです。2年前期と後期でどのくらいループリックの得点が変わったかを示しています。この評価を行ってとくに興味深かったのは、学生の感想です。例えば「ステップⅢを行うことで、PBLが将来現場に出た時に役立つものになるのだとわかりました」という声がありました。

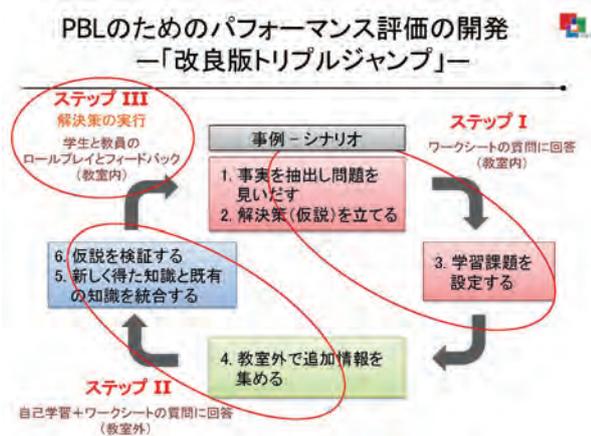


図18

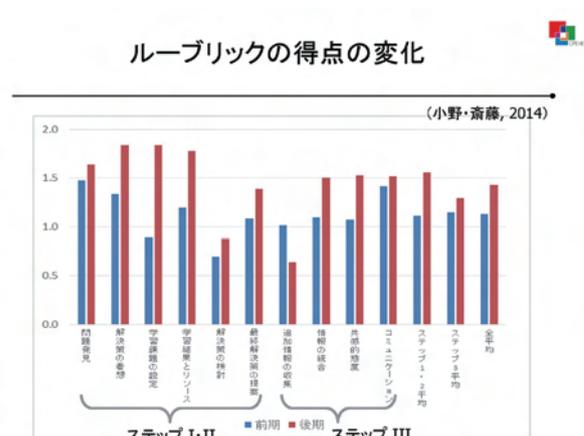


図19

PBLのシナリオは教員が実際に出会った症例などを使って作られているので、将来出会うかもしれない問題だということを学生は理解しているはずだと思っていたのですが、実はそうでもなかったようです。今回、先生を模擬患者役にしてロールプレイを行ってもらったのですが、自分で一人で考え、実際にやってみて、初めてPBLが何のために行われていたのかがよくわかったということなのですね。思った以上の手ごたえでした。

## 6. 学習評価のこれから

最後に学習評価のこれからについてお話しします。文科省の調査によると、ルーブリックは実は思ったほど広がっていないのです。よくルーブリックの効用として、「ぶれない」とか「時短」とか言われますが、実際は、ルーブリックを使ってもそれなりにぶれます。また、時短についても、アメリカのように教員が学生にきちんとフィードバックしている場合は、コメントする代わりにルーブリックを素早くチェックすればよくなるので時短になりますが、日本の場合、大抵はレポートを出しっぱなしで終わりということが多くですね。それだと、これまでよりも評価が面倒になります。そのため、アメリカのようにTAをレポート評価で活用できず、また、これまでほとんどフィードバックをしてこなかった日本の教員にとっては、必ずしも時短になるわけでもありません。したがって、日本の場合、こうした面でのルーブリックの効用は小さいのではないかと思います。むしろ、一番大きいと思うのは、評価主体としての学生を育てるところです。

先ほどのPBLのように、評価自体が学習になるような場合、ルーブリックは有効です。学習と評価を結びつける言葉として、「学習の評価」、「学習のための評価」、「学習としての評価」という言葉がありますが、3番目の「学習としての評価」は、学生自身も評価主体になる評価です。特に今回行ってみて痛感したのは、

評価そのものが学習経験になるということでした。評価を評価としてのみ行うのではなく、学習経験にもなるような評価を作るということです。そのためにも、評価それ自体が学びになるような評価をいかにデザインするかが重要です。そうなれば、パフォーマンス評価、あるいはその中で用いられるルーブリックにも意味があるのではないかと、思います。

最後に、さまざまなタイプの学習評価の関係についてお話ししておきたいと思います。現在、大学教育に対して、アカウントビリティや質保証の要請が強くなっています。その中で、タイプⅡでは、IRのための学生調査が普及してきました。タイプⅢでは、OECD-AHELOや業者作成による標準テストが実施されるようになってきています。私は、タイプⅣにこだわっているのですが、それは学習としての評価という可能性がタイプⅣに最もあると思っているからです。

機関レベルでのルーブリックによる評価は、アメリカでは2013年の調査で約70%に達し、最近の数年間で大きく伸びていますが、日本はたった約3%です。やはり、日本の場合さまざまな障害があると思います。特に評価負担の大きさに繋がらないようにしなければ、おそらくあまり広がっていかないのではないのでしょうか。

## 7. 今後の課題

最後に今後の課題を挙げておきたいと思います。1番目は、学生に学問分野に根ざしながら、一般性(汎用性)を持つような能力をどのようにして形成していくかということです。今日は、新潟大学歯学部を取組を引きながら、問題解決能力の例をお見せしたつもりです。2番目には、それぞれ4つのタイプの学習評価の特徴を把握して、その目的に応じて使い分けたり、組み合わせたりすることを通して、学習成果を包括的に評価していくということです。このときにはおそらく、政策レベルと実践レベルの違いやプログラムレベル(機関レベル)や科目レベルの

違い、そのようなことも考慮しなければならないだろうと思います。3番目には、評価疲れに陥らないようにするためにはどうすればよいかということです。私は、一石二鳥を狙うことが大切だと思っています。評価を通じて学びを深めるというのが一石二鳥ですが、評価開発を通じてFDが進んでいくこともありますので、一石二鳥になるだけではなく、三鳥でも四鳥でも狙っていけるとと思います。ちょっとした知識などは、パフォーマンス評価などする必要はないので、より評価負担の軽い評価と組み合わせを使って、ここぞという時にしか重たい評価を使わないという、そういうメリハリも必要ではないかと思っています。

きょうは長時間、ご清聴ありがとうございました。

#### 《質疑応答》

質問：

制度設定のところで気を付けることについてお伺いします。特に何か先生がされた点で、もう少し参考になるようなことを教えていただければと思います。可能な限りで結構ですのでよろしくお願いします。

松下先生：

例えば、PBLの評価があります。普段のPBLはグループで行っています。グループで協働的に学ぶことにはいい面もたくさんあります。しかし、逆にグループ学習では、フリーライダーの問題が生じたり、責任の所在が明確にならなかったりすることもあります。それで、PBLの評価では課題に一人で取り組んでもらいました。一人PBLみたいなものです。一人ですが一週間は空けてありますので、その間に資料を調べたり、人と相談したりすることはできます。そして最後のところでは、自分が学んできたことを、実際にパフォーマンスさせます。これはとても重要です。

私は今、大阪府の教育センター附属高校で演

劇ワークショップのパフォーマンス評価を行っています。そこでは、クラスの中で7～8人ずつのグループを組んで、ゼロから自分たちで演劇を作っていきます。平田オリザさんの流れを汲む演劇ワークショップですが、自分たちで演劇を作っていくって、まずクラスの中で演じ、その後で今度は各クラス代表になったグループがホールの舞台上で演じます。そういうパフォーマンスがあるごとに生徒たちはどんどん伸びていきます。こうした機会を設けるのはとても重要だし、単に教室に閉ざされないで、他の人からも評価をしてもらう機会を設けるという点で重要ではないかと思っています。

#### 文献

AAC&U (2007). College learning for the new global century: A report from the National Leadership Council for Liberal Education & America's Promise.

Alverno College Faculty (1994). Student assessment-as-learning at Alverno College. Alverno College Institute.

コテ, J. (2014). 「アイデンティティ資本モデル—後期近代への機能的適応—」溝上慎一・松下佳代(編)『大学・高校から仕事へのトランジション—変容する能力・アイデンティティと教育—』ナカニシヤ出版.

Earl, L. M. (2003). Assessment as learning: Using classroom assessment to maximize student learning. Thousand Oaks, CA: Corwin Press.

濱名篤他(2012). 「体系的なカリキュラム構築と学習成果の可視化のためのルーブリックの構築・活用」大学教育学会第34回大会ラウンドテーブル.

Gipps, C. V. (1994). Beyond testing: Toward a theory of educational assessment. 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.
- Gonzalez J., & Wagenaar, R. (2008). Tuning educational structures in Europe - Universities' contribution to the Bologna Process: An introduction (2nd ed.). ゴンサレス, 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—』(深堀聡子・竹中亨訳) 明石書店.
- 松下佳代(2007). 『パフォーマンス評価』日本標準.
- 松下佳代(2010). 「〈新しい能力〉概念と教育—その背景と系譜—」松下編『〈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.
- 松下佳代(2012). 「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—」『京都大学高等教育研究』第18号, 75-114.
- 松下佳代・小野和宏・高橋雄介(2013). 「レポート評価におけるルーブリックの開発とその信頼性の検討」『大学教育学会誌』第35巻第1号, 107-115.
- 松下佳代(2014). 「学習成果としての能力とその評価—ルーブリックを用いた評価の可能性と課題—」『名古屋高等教育研究』第14号, 209-229.
- 松下佳代・京都大学高等教育研究開発推進センター編(2015). 『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.
- 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巻1号, 123-132.
- 小野和宏・大内章嗣・前田健康(2011). 「学習者主体PBLカリキュラムの構築—新潟大学歯学部口腔生命福祉学科7年のあゆみ—」『新潟歯学会誌』41(1), 1-12.
- McTighe, J., & Wiggins, G. (2004). Understanding by design: Professional development workbook. Alexandria, VA: Association for Supervision and Curriculum Development.
- Palomba, C., & Banta, T. (1999). Assessment essentials: Planning, implementing, and improving assessment in higher education. San Francisco: Jossey Bass.
- Rhodes, T. (Ed.) (2010). Assessing outcomes and improving achievement: Tips and tools for using rubrics. Washington, DC: Association of American Colleges and Universities.
- Rhodes, T. (June 29, 2013). Effectiveness of Liberal Education and America's Promise (LEAP) and VALUE Rubrics. (講演資料)
- Rhodes, T. L., & Finley, A. (2013). Using the VALUE rubrics for improvement of learning and authentic assessment. Washington, DC: Association of American Colleges and Universities.
- 田中耕治編著(2005). 『よくわかる教育評価』ミネルヴァ書房.
- 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—』ぎょうせい.
- 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.
- 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.