

平成27(2015)年度FDセミナー

# 能動的学びの実践と 学習成果の評価

予測困難な現代社会において、大学は、学生に専門的知識を修得させるだけでなく、生涯にわたって有用となる能力や学習姿勢を身につけさせることが求められています。

そこで、平成27年度FDセミナーでは、能動的学び（アクティブ・ラーニング）を通じて身についた“学習成果の評価方法”に着目し、京都大学の松下教授から「ディープ・アクティブラーニングの方法とその評価」と題したご講演をいただきました。

また、本学教員による「能動的学びを促す授業の実践事例」の紹介やパネルディスカッションを実施しました。

## ○基調講演

「ディープ・アクティブラーニングの方法とその評価」

講師：松下 佳代 氏

（京都大学高等教育研究開発推進センター 教授）

11/20

15:00～  
17:50

## ○学内事例報告

「能動的学びを促す授業の実践事例」

報告者：福田 公子（都市教養学部 理工学系 准教授）

西内 信之（システムデザイン学部 准教授）

荒戸 寛樹（都市教養学部 経営学系 准教授）

## ○パネルディスカッション

パネリスト：松下 佳代、福田 公子、西内 信之、荒戸 寛樹

大森 不二雄（大学教育センター 教授）

司 会：松田 岳士（大学教育センター 教授）



# ディープ・アクティブラーニングの方法とその評価



まつした かよ

1991年京都大学大学院教育学研究科博士後期課程研究指導認定退学。博士（教育学）。京都大学教育学部助手、群馬大学教育学部助教授、京都大学高等教育教授システム開発センター助教授を経て、2004年より現職。専門は教育方法学、大学教育学。特に能力、学習、評価をテーマに研究と実践を行っている。主な著作に、『パフォーマンス評価』（日本標準、2007）、『〈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 -』（編著、ミネルヴァ書房、2010）、『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 -』（編著、勁草書房、2015）などがある。

京都大学高等教育研究開発推進センター 教授 **松下 佳代**

大学教育における「教授から学習への転換」の鍵として、近年アクティブラーニングが注目されている。今回の講演では、「深い学び」につながるアクティブラーニングの方法を、京都大学の松下教授が提案する。また、そのような学習を通じて身についた学習成果の評価方法（パフォーマンス評価、ポートフォリオ評価、ルーブリック等）についても考える。

## 1. はじめに

ディープ・アクティブラーニングという言葉は、私の造語です。「生徒・学生が他者と関わりながら、対象世界を深く学び、自分のこれまでの知識や経験と結びつけると同時に、これからの人生につなげていけるような学習」という意味で使っています。

本日は、学習をアクティブ（能動的学習）だけでなく、ディープ（深い学習）とも重ね合わせて考え、そのような学習では、何を、どのように評価するのかをお話していきたいと思えます。

アクティブラーニングという言葉は、アメリカの高等教育界で1980年代ごろから使われ始めました。その議論を総括したのがBonwell & Eisonの論文です。彼らはアクティブラーニングを、「学生にある物事を行わせ、その行っている物事について考えさせること」と定義しています。

日本の場合は2011年ごろからニュースのヘッドラインでアクティブラーニングが取り上げられるようになり、2014年後半ごろから、さらに話題にのぼるようになりました。

アクティブラーニングが普及した背景には、世界的に見ると教授（teaching）から学習（learning）への焦点の移動があります。また、インターネット時代において、学校・大学の教育・学習空間としての価値が問い直されていることも挙げられます。私の所属する京都大学の高等教育研究開発推進センターでも、大規模公開オンライン講義（MOOC）の開発などを行い、京大の先生方の授業を世界に発信しています。

特に日本において、昨年冬ごろから普及してきた背景には、明らかに「新しい時代に必要となる資質・能力の育成」のための“公定の教育方法”としての政策

的な後押しがあります。

## 2. 教育政策の中のアクティブラーニング

今の教育政策には、大きく二つの特徴があります。

一つは、小学校から大学までの「一体的な改革」です。アクティブラーニングという言葉が特に大きく取り上げられたのは2012年8月の質的転換答申からで、このときは大学でのアクティブラーニングが言われていました。そして、2014年11月、現在改訂作業が行われている次期学習指導要領について、文部科学大臣が中教審に対して諮問した中で、アクティブラーニングという言葉が多く使われました。つまり、小学校から高校までのアクティブラーニングがここで出てきたのです。そして、翌12月の高大接続答申の中でも、アクティブラーニングが高校教育と大学をつなぐキーワードとして使われています。

もう一つの特徴は、目標と方法と評価の「整合性（alignment）」です。今の教育政策では、「資質・能力」や「学力の三要素」、「生きる力」といった言葉が目標として掲げられています。これらは、これまで小学校から高校について言われてきたことですが、高大接続答申では、大学にも当てはめられています。

そして、それを達成する方法が「アクティブラーニング」であるとされています。質的転換答申では「学修者の能動的な学修への参加を取り入れた教授・学習法の総称」とかなり広く定義されていましたが、次期学習指導要領や高大接続答申の議論では、「課題の発見と解決に向けた主体的・協働的な学習・指導方法」という定義が使われるようになっていきます。

さらに、今の教育政策では、そのための評価方法として、大学におけるパフォーマンス評価、ポートフォ



リオ評価、ルーブリックなどを挙げ、各大学における個別の入学選抜においても、それらの評価方法と似た方法を使うように、としています。目標と方法と評価が非常に強いつながりを持ったかたちで提案されていることが、大きな特徴と言えるでしょう。

### 3. アクティブ・ラーニングに対する危惧

現在、アクティブ・ラーニングは、初等・中等教育では大学以上に大合唱という感じになっていて、見ていると少し危惧されるところがあります。アメリカのカリキュラムや評価の研究者である Wiggins & McTighe は、「教授デザインにおける双子の過ち」ということを言っています。一つは教科書などの内容の網羅に焦点を合わせた指導（網羅主義）、もう一つは活動に焦点を合わせた指導（活動主義）です。元々、日本の場合、大学の講義一辺倒の授業では駄目だという網羅主義に対する批判として、アクティブ・ラーニングが出てきたと言えると思います。日本の教育方法史を振り返ってみると、絶えず対立する二つの項目のあいだで、振り子のように揺れる動きを繰り返してきました。今回でいえば、アクティブ・ラーニング vs. 講義、コンピテンシー・ベース（能力） vs. コンテンツ・ベース（知識）というかたちになっています。しかし、二項対立として捉えるのではなく、どうすれば両立できるかを考える必要があるのではないのでしょうか。

#### アクティブ・ラーニングに対する危惧

- 教授デザインにおける双子の過ち (Wiggins & McTighe, 2005)
  - 内容の網羅に焦点を合わせた指導 = 網羅主義
  - 活動に焦点を合わせた指導 = 活動主義
- 二項対立的な捉え方の中で、再び振り子の動きが繰り返されるおそれ
  - アクティブ・ラーニング vs. 講義
  - コンピテンシー・ベース（能力） vs. コンテンツ・ベース（知識）

二者択一ではなく、両者を両立させるには？

9

### 4. ディープ・アクティブラーニングの提案

そのような観点から提案しているのが、ディープ・アクティブラーニングです。望ましい学習の在り方を考えたときに、アクティブ・ラーニング一辺倒になるのは不健全だと私は考えています。アクティブ・ラーニングは学習者の主体的・能動的な学習への参加を求めている、確かにそのこと自体はとても重要だと思

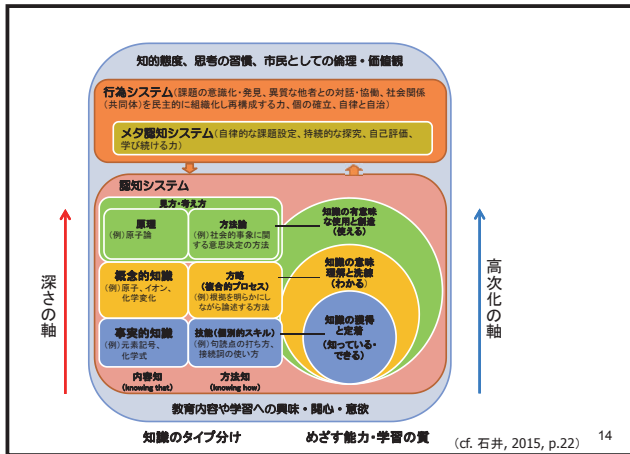
ますが、望ましい学習の在り方がこれで全て語り尽くせるわけではありません。ディープ・ラーニングは学習内容や質の深さに注目するもので、既に 1970 年代ごろから言われてきているのですが、これもまた学習の望ましい在り方だと思うのです。ですから、私はアクティブ・ラーニングをディープ・ラーニングとうまく組み合わせるやり方を考えていきたいと思っています。

以前、ある大学の FD セミナーに講演者として呼ばれたときのことで。その大学ではずっと教職員と学生が対話形式で FD を行っていて、これまでの議論の内容を事前に資料として送っていただきました。その資料には、学生からは「授業は盛り上がったが、果たして学生が深いところまで内容を理解して盛り上がっていたのかどうかは分からない」、教職員からは「活動的に見えても実は深く考えていない授業もあり、教育の質という面ではディープに学べるような理論や仕掛けがある授業が良い授業と言えるかもしれない」といった意見が書かれていました。学生と教職員で立場は違いますが、かなり似たような問題意識を持って議論されていることが分かります。そのようなことを踏まえても、この二つをうまく結びつけていくことが必要ではないかと考え、ディープ・アクティブラーニングという言葉を作ったわけです。

そう考えて改めて調べてみると、最近、特にアメリカで Deep (Deeper) Learning が非常に注目されていることに気がきました。全米研究評議会 (National Research Council : NRC) は、2012 年に出した報告書の中で「21 世紀型コンピテンス」と「Deeper learning」という考え方を提案しています。「21 世紀型スキル」という概念は日本でも知られていますが、スキルだけでいいのかという問題意識が、「21 世紀型コンピテンス」という言葉には込められています。スキルだけでなく、内容領域の深いところにある知識も教えないと、必要な場面で使いこなす能力にはならないということです。そして、今まで自分が出会ったことがないような場面や課題に直面したときに使える知識を身につけるための深い学習を、Deeper learning と言っています。

今日、特にお話ししたいのは、この「深さ」と、より高次の能力を結びつけることの重要性についてです。より高次の能力とは、単に記憶するだけではなく、新しい課題に応用したり、分析したり、総合したりできる力のことで。

京都大学の石井英真氏は、「知識のタイプ」と「めざす能力・学習の質」の関係を次のような図にまとめています。



注目すべきは、深さの軸が示されていることです。まず知識は、内容知 (knowing that) と方法知 (knowing how) に大きく分けることができます。そして、一番浅いところにある知識には、事実に基づく知識や技能 (個別的技能) があります。一番深いところにあるのが原理や方法論で、その間をつなぐのが概念的知識や方法論です。このように、いくつかの知の構造を階層的に捉えることができます。

今はインターネット時代で、知識を学校で教える必要はあまりなく、検索すればすぐに出てくるとよく言われますが、それは浅い知識です。その事実に基づく知識をもっと深くにある概念的知識や原理と結びつけて理解させることは、一人でネット検索しているだけではなかなか身につけません。仲間と議論したり、先生から指導を受けたりして、より深い知識を身につけていくことができます。これが深さの軸です。

もう一つは、高次化の軸です。単に知っていたり、できたりするだけでなく、その意味を理解して自分で洗練させていくことと、いろいろな場面で使って必要に応じて新しく創造していくことの二つを結びつけることが、ディープ・アクティブラーニングでは必要になります。

評価をするときにも、どの部分の知識を見ようとしているのか意識する必要があります。例えば、客観的テストで正誤を○×で答えさせるのは、浅い知識を見る場合にはいいのですが、もう少し深い知識を場面の必要性に応じて使いこなせるかどうかは、客観的テストだけではなかなか見ることができません。

## 5. 学習評価の構図

さて、ここからもう少し詳しく評価のお話をしてみましょう。評価についても、私は大きく二つの軸で捉えられると思っています。一つは直接評価と間接評価で、もう一つは量的評価と質的評価です。この二つ

の掛け合わせにより、大学・学校で現在行われているいろいろな評価のタイプを捉えることができます。

間接評価とは、学生が「何ができているか」、つまり学生の学習行動や学習についての自己認識を通じて、学習成果を間接的に評価することです。一方、直接評価は、学生が「何ができるか」、つまり学生の知識や技能を実際に出させて、それを直接的に評価することです。

### 間接評価と直接評価

(Palomba & Banta, 1999)

#### ● 間接評価

- 学生の学習行動や学習についての自己認識を通じて、

学習成果を間接的に評価 **何ができているか?**

- 質問紙による「学生調査」

- 【例】 NSSE (National Survey of Student Engagement) JCIRP (日本版大学生研究調査プログラム) (山田, 2012)

#### ● 直接評価

- 学生の知識や技能の表出を通じて、

学習成果を直接的に評価 **何ができるか?**

22

もう一つの軸である量的評価と質的評価は、おなじみだと思います。例えば、同じ大学で3年前と現在を比較したり、他の大学や学部・学科間で比較したりするときには、質的評価ではなかなか比較しにくいので、量的評価が使われます。選抜するときや外部に向けて説明する際にも量的評価が客観的、簡潔で分かりやすいという理由で使われます。一方、目の前の学生に何か指導したり、改善したりするときには、質的評価が非常に有効です。

### 量的評価と質的評価

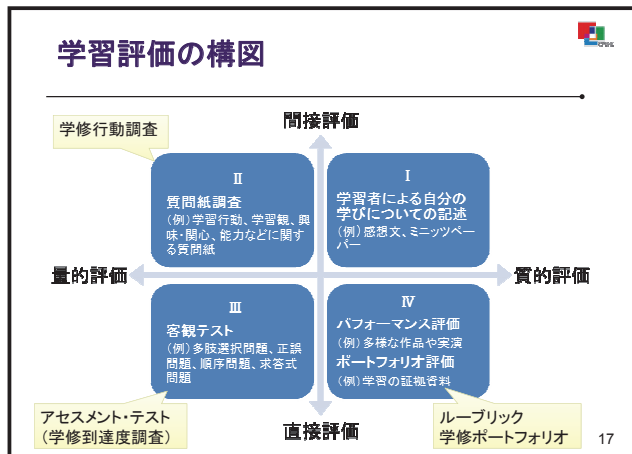
	量的評価	質的評価
評価データ	量的データ	質的データ
評価対象	集団または個人	個人
評価目的	比較、選抜、アカウンタビリティなど	学習や指導の改善など
評価課題	細かく分割された問題 文脈独立的	複合的な課題 文脈依存的
評価基準	客観性を重視	間主観性を重視
評価結果	数値	文章や数値
評価機能	主に総括的評価	主に形成的評価
評価方法	客観テスト、質問紙調査など	パフォーマンス評価・真正の評価・ポートフォリオ評価、感想文など

23

ただ、質的評価でよく問題になるのが、どうしても主観的になりやすいことです。主観的になることは止めようがないと思いますが、恣意的、独断的にならないようにする必要があります。そのためのいろいろな道具立てについては、後ほどお話ししたいと思います。

質的転換答申では、各大学はアドミッション・ポリシー、カリキュラム・ポリシー、ディプロマ・ポリシーの三つに加えて、アセスメント・ポリシーを持つべきであり、特に学習成果の評価に当たっては、学習時間の把握といった「学修行動調査」や「アセスメント・テスト（学修到達度調査）」、「ルーブリック、学修ポートフォリオ」等、どのような測定手法を用いたかを併せて明確にしなければならないと言っています。

ただし、これらの評価方法は、それぞれかなり性格が異なるものです。縦軸に間接評価と直接評価、横軸に量的評価と質的評価をとって学習評価の構図を描くと、学修行動調査、アセスメント・テスト、ルーブリック、学修ポートフォリオは、それぞれ異なる象限に分類されます。つまり、それぞれ異なる性格を持った評価なのです。



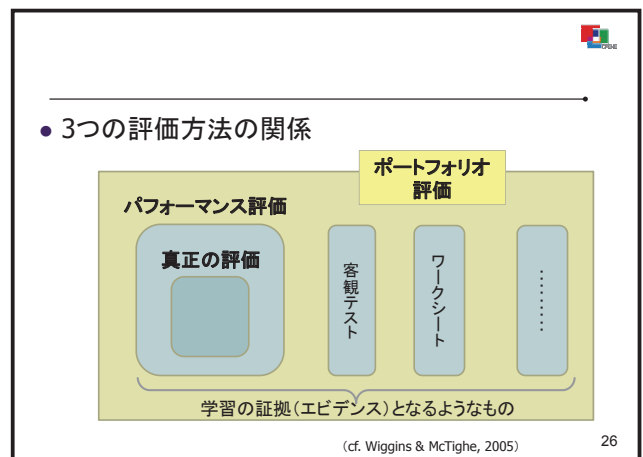
例えば、毎回の授業の最後に、今日学んだことを書かせている先生がいらっしゃるかもしれません。ミニツツペーパーと言われます。1分で書けるぐらいの、振り返りのためのペーパーです。これは学生自身が書いた自分の学びについての記述を質的に評価するものなので、タイプIです。また、質問紙調査で、学生に自分がどんな能力が身についているか、どんな知識があるかを、5件法などを使って答えさせるものがあります。これは量的に尋ねて学生自身の自己認識を通じて学習成果を評価していくので、タイプIIになります。それから、皆さんおなじみのセンター試験等の客観テストは、量的に直接評価しますから、タイプIIIになります。そして、大学で非常によく使われているレポートや卒業論文、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、質的に直接評価するので、タイプIVになります。

このようなかたちでの学習成果の評価は、特に大学では重要なのですが、皆さんは自分も学生としてその

ようなものを受けてきたという経験値、あるいは専門分野について学んだことを基に評価されていると思います。それを、もう少し自覚的に行おうという提案をしたいと思います。

今日、特にお話しするのは、タイプIVに分類される評価です。このタイプの評価法は、これまでの客観テストでは、学生・生徒が身につけていることをきちんと評価できていないということによって生まれた評価法です。

その一つであるパフォーマンス評価は、実際に何かをさせてみて行う評価です。作品や実演を例として挙げていますが、レポートや卒業論文なども学生の作品、模擬授業なども実演と捉えてください。一定の意味ある文脈（課題や場面など）の中で行われる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・スキルを総合的に活用する能力を評価するのがパフォーマンス評価です。



パフォーマンス評価の中でも、特に課題の本物らしさ（真正性）にこだわった評価が真正の評価です。例えば、自動車学校の試験にはタイプIIIのペーパーテストもありますが、仮免許試験では自動車学校の敷地内を実際に運転させます。それがパフォーマンス評価です。しかし、本試験になると路上に出て、より実際の状況に近い場面で運転させます。これが真正の評価です。

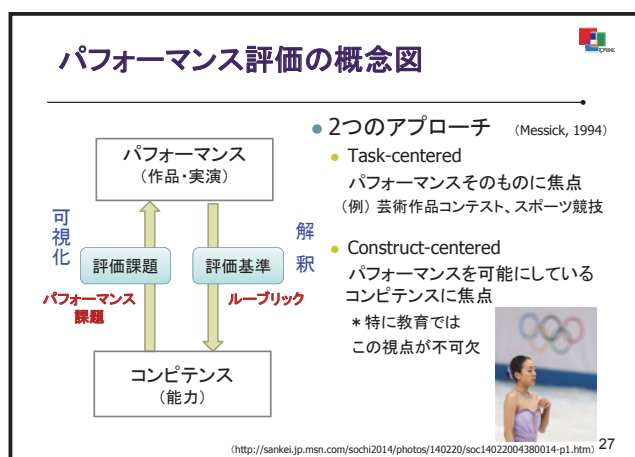
それに加え、ポートフォリオ評価があります。ポートフォリオとは、もともとは、画学生などが自分の作品を画商に売り込みにいくとき、自分の作品の中から自信のあるものを集めて持って行くための紙ばさみのことです。学習の中でレポートを書いたり、絵を描いたり、テストを受けたりといった証拠資料のようなもの（エビデンス）を束ねて学びのプロセスを評価するのが、ポートフォリオ評価です。

ここでは、特にパフォーマンス評価についてお話しします。「コンピテンス（能力）」は、どんなもので



あってもそれ自身を観察することは不可能です。私が聞いたところでは、各地に力石というものが残されているそうです。その石をきちんと持ち上げることができたら、一人前の男になったことを示すという石です。その力石と同じように、私たちは観察不可能な能力をいったん何らかのかたちで可視化し、パフォーマンスにして、そのパフォーマンスを解釈することで能力を判断しています。そのときに特に重要なのが、評価課題と評価基準です。ルーブリックは、評価基準の一形式です。

パフォーマンス評価には二つのアプローチがあります。一つは、パフォーマンスそのものを見ること (Task-centered) です。例えば芸術作品のコンテストやスポーツ競技では、その場のパフォーマンスだけが評価の対象になります。もう一つは、パフォーマンスを可能にしているコンピテンスも併せて見ていくこと (Construct-centered) です。Construct とは、観察される現象の背後にあるものを説明するときに使われる構成概念という意味です。特に教育では、この視点が不可欠です。



パフォーマンス課題 (performance task) とは、学習者のパフォーマンスを評価するためにデザインされた、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を言います。大学教育では、パフォーマンス課題が本当にたくさん使われています。医療系や教員養成などの実技や、芸術分野の演奏・作品、もっと一般的に言うとレポート、卒業論文、プレゼンなども全てパフォーマンス課題になります。しかし、その評価基準は、これまでほとんど主観に委ねられていました。

それを目に見えるかたちにしたのが、ルーブリックです。能力やパフォーマンスは、とても優れたものから全然駄目なものまでがつながっている連続体です。その評価に当たっては、パフォーマンスやその背後に

ある能力の質の差異で区切っていくとともに、どういう多面性を持っているかも見ていく必要があります。

ルーブリックは、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、いくつかの観点 (規準) とレベルを設定して質的な差異を区切り、例えば観点4・レベル1はどのような状態のものをいうかを言葉で書き表します。要するに、教師 (専門家) の鑑識眼を明示化したものであり、それによって教員とティーチング・アシスタント (TA)、学生が評価基準を共有することができるようにするツールと言えます。また、パフォーマンスの質を量的表現に変換する働きもあります。

### 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
  - 「質的転換答申」で言及
  - アメリカでは70%、日本では初等中等教育から普及

	レベル4	レベル3	レベル2	レベル1
観点1 (規準)	.....	.....	.....	.....
観点2	.....	.....	.....	.....
観点3	.....	.....	.....	.....
観点4	.....	.....	.....	.....

29

ルーブリックは質的転換答申の中で学習成果の測定方法として言及されましたが、日本では、大学よりもむしろ初等・中等教育から普及し始めています。一方、アメリカでは既に70%ほどの大学で使われています (2013年のデータ)。アメリカで評価方法として最も使われているのは全米規模で行われている学生調査ですが、その次にルーブリックやクラス単位で行われるパフォーマンス評価などが挙げられています。パフォーマンス評価でルーブリックを使うことにより、質的な評価だけでなく量的な評価にもなり、統計的な分析などもかけられるようになるからでしょう。

## 6. 評価の具体的事例

### 6-1. 修士論文の評価

いくつか具体例をお示ししたいと思います。まず、愛媛大学教育学研究科の修士論文の評価の事例です。ネット上に掲載されていたものから、一部を抜粋させていただきました。ルーブリックは4観点でA~Dの4レベルの評価があり、修士論文を指導教員とともによりよいものに完成させること (各観点でB以上になった状態で修士論文を提出すること) が目標です。

● 修士論文用リサーチルーブリック

領域	評価項目	A	B	C	D	評価理由
テーマ設定	現代の教育課題への貢献	B	現代社会あるいは地域における教育の課題の解決や理解の深化に関連するテーマを設定している			
	発展可能性	B	より重要な研究へと発展することが可能なテーマである			
研究計画	計画・準備	B	指導教員との協議を通して計画書を作成し、研究シナリオ、データ収集、分析、執筆など具体的な活動のスケジュールを明確にしている			
	研究倫理	B	研究に関わる倫理上の問題について、十分な考慮と必要な対応を行っている。研究活動を行っている			
研究の目的	目的の明示	B	研究の目的は述べられており、その目的を達成するためのどのように進めていくのかもほぼ明らかである			
	研究方法の妥当性	B	研究目的を達成するために適していると考えられる研究方法を採用している			
データの扱い	記述法・ルール	B	論文の本文は学術的な記述法で書かれ、当該分野の学会で一般的に利用されている執筆規則にも準拠している			
	データの量	B	研究目的を達成するために選択した研究方法、分析方法を実施するのにほぼ十分な量のデータ・資料を収集している			
結果の表現	分析手法	B	研究目的を達成するために選択した研究方法にふさわしい分析手法が行われており、当該分野における一定の水準に到達している			
	結果の表現	B	結果を適切に表現するために必要な図表等がおおよそ作成されており、ほぼ問題なく配置されている			
結果の解釈とまとめ	結果の解釈	B	参考文献や得られたデータに基づいて適切な公平な解釈をおこなっている。予想や仮説に一致しない結果は例外として処理している			
	結果のまとめ	B	当該分野において有意義な知見や発見を参考文献や得られたデータに基づいて提供している			
成果	成果の公表	B	関連する学会あるいは学内の発表会で研究成果を十分に発表することができた			
	実践への貢献	B	少し手を加えれば、研究によって明らかとなった知見や成果物を教育現場等の実践者に役立ててもらえる			

(http://www.ed.ehimeu.ac.jp/~rubric/index.html) 34

修士課程に入ったとき、ルーブリックの意義と目的、使用法について説明し、半期に1度、院生と指導教員それぞれが個別にルーブリックに従って評価を行い、それを突き合わせて評価結果と評価理由について協議し、改善点と課題を明確にしています。それを4回実施し、つなげることによって、その学生の成長が把握できるというポートフォリオ的な使い方もできるようになっています。そして、最後に実際の論文が評価対象になっているということです。

### 6-2. コンセプトマップを使った評価

次に、私が直接関わった、京都大学文学部でのコンセプトマップを使った評価です。京都大学の哲学系入門科目では、5人の講師によるリレー講義を行っています。その授業の評価方法は、5人の講師のテーマから一つを選んでレポートを書くというものでした。この方法は、深く掘り下げてレポートを書くという意義はあるのですが、五つのテーマ全体から何を学んでいるのかは分かりません。

そこで、テーマ全体から何を学んでいるのかを評価する方法として、コンセプトマップを取り入れました。5人の講師が半期のリレー講義で扱った複数の異なるテーマやそのキーコンセプトについて学んできたことを、「哲学的に考えると」という当該分野の本質的な問いに答えるかたちで、学生自身が結びつけていくことを、最後の授業で行いました。2~3人ずつのグループを作り、コンセプトマップをそれぞれ発表して、最後に先生が講評します。この評価に関して、担当講師の合議によりルーブリックを作成してみました。コンセプトが正しく理解されているか、異なるテーマ同士を結びつけるためにコンセプトを新たに作っているか、コンセプトをつなぐリンクが適切か、「哲学的に考える

とは」というテーマときちんと結びつけられているかといった観点で、0~3の4段階で評価しました。



ルーブリックなしで評価するよりも評価者間の信頼性が高く、評価負担も大きくないという評価面での意義もさることながら、学習面でも大きな意義がありました。学生は「15回の講義を追体験できた」(リフレクション)、「ペアで作るので、自分にはない発想が出てきた」(協働)、「バラバラに思えた五つの授業が一つにまとまった」(知識の統合)といった意義を見出すことができ、教員からも「物事を突き詰めると別の話につながっていくということを体験するのは、哲学としてよかったと思う」といった意見が出ました。このようなことから、それ自身が学習にもなるような評価をデザインすることの重要性を再認識しました。

ルーブリック

規準	規準の説明	3 excellent	2 good	1 satisfactory	0 unsatisfactory
コンセプトの理解	授業に関連したコンセプトを理解しているか。	各授業で扱われたコンセプトが豊富に含まれており、その使い方が適切である。	提示されたコンセプトはほぼ適切に用いているが、それ以外の各授業で扱われたコンセプトはあまり用いていない。	コンセプトがかなり含まれていない。	レベルを満たさないもの
コンセプトの抽出	授業間を関連づけるために、新たなコンセプトを創出し、それを効果的に用いているか。	授業間を関連づけるために、独自のコンセプトを創出し、それを効果的に用いている。	授業間を関連づけるための独自のコンセプトを創出しているが、効果的とはいえない。	授業間を関連づけるための独自のコンセプトがほとんど含まれていない。	レベルを満たさないもの
リンクの構成	コンセプト間の関係を適切に結んでいるか。	適切なリンクが数多く張られている。階層性や分岐構造が明確に分けられる。複雑な適切なクロスリンクがある。	適切なリンクが張られているが、数が多くない。階層性や分岐構造が不十分である。クロスリンクが一部に留まっている。	リンクが張られているが、不適切である。階層性や分岐構造があまり見られない。クロスリンクがある授業のテーマと別の授業のテーマをつなぐリンク。	レベルを満たさないもの
リンク語の適切性	コンセプト間の関係を適切な語で表現しているか。	適切なリンク語が数多く含まれており、それによってリンクの意味が明確になっている。	リンク語がかなり含まれているが、リンクの意味が明確ではないものが多い。	リンク語が少なく、または不適切なリンク語がかなり含まれている。	レベルを満たさないもの
中心テーマとの関連性	中心テーマに即して、授業内容を関連づけているか。	授業内容を、中心テーマに関連づけて、その関連性を一貫性をもってコンセプトマップ全体で表現している。	中心テーマに即して授業内容を関連づけているが、関連づけが部分的である。	授業内容を関連づけているが、中心テーマとずれている。	レベルを満たさないもの

・ 担当講師の合議によりルーブリックを作成  
・ ルーブリックなしで評価するより高い評価者間信頼性、評価負担も大きくない

40

### 6-3. PBL の評価

次に、私が共同研究を行ってきた新潟大学歯学部の新潟大学歯学部のPBL (Problem-Based Learning) の評価についてお話しします。PBLとは、問題解決能力や自己学習能力、知識の統合的な理解などを身につけさせるために行われている学びの方法です。新潟大学歯学部では、先生方が実際に出会った症例をシナリオ化し、2004年度からPBL テュートリアル教育を実施しています。これは、学生7

## 7. まとめと今後の課題

最後に、ここまでお話ししてきたことを、前にお話しした、「めざす能力・学習の質」と関連付けてまとめたいと思います。「知識の意味理解と洗練（わかる）」や「知識の有意味な使用と創造（使える）」というところを見るには、コンセプトマップやPBLのトリプルジャンプのようなかたちでのパフォーマンス評価が有効です。ポートフォリオには、学習者自身が作る評価という意味合いもあります。

改めて申し上げたいのは、「学習としての評価」という考え方です。学習と評価を結びつける言葉として、「学習の評価（成績を決定する／教師が評価主体）」「学習のための評価（授業や学習改善に結びつく／教師が評価主体）」がありますが、それに加えて「学習としての評価」があります。学生自身も評価主体となり、評価でありながら学習経験にもなるような評価のこトです。今日ご紹介した三つの評価の事例は、いずれも評価自体が学びになるような「学習としての評価」の例でした。

アクティブ・ラーニングの評価として、最近、質問紙調査がよく使われるようになってきていますが、私は、タイプⅣの評価をタイプⅡの質問紙調査で代替することはできないと考えています。能力が低い人はメタ認知能力の低さゆえに、自分自身の能力の低さに気付かないだけでなく、自分の能力を過大評価する傾向があり、逆に能力が高い人は、自分ができるとは他の人もできるのではないかと想定するため、割と低く評価する傾向があるからです。これは、「ダニング＝クルーガー効果」と呼ばれています。

私たちが取ったレポート評価のデータによれば、アカデミック・ライティングについての学生の自己評価と教員の評価にはほとんど相関がありませんでした。しかし、学生の自己評価と質問紙の回答の間にはかなり相関があります。そして、教員の評価と学生の質問紙は全く無相関で、マイナスの相関の項目すらあります。これを見ても、直接評価を間接評価で代替するのは難しいことがわかります。

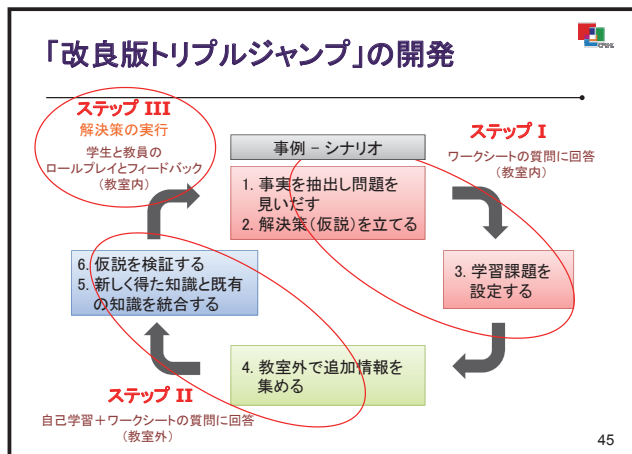
今日は、四つのタイプの評価についてお話ししました。それぞれに良さがありますが、今後の課題はその使い分けと組み合わせです。ただ評価が増えれば評価疲れが心配されるので、一石二鳥をねらって、単なる評価ではなく、それが学習にもつながる評価をデザインしていくことと、より負担の軽い評価と重い評価をうまく組み合わせることが必要だと考えています。

長時間、ご静聴いただき、ありがとうございました。

～8人のグループに教員がチューターとして入り、その支援の下でPBLを行うというものです。まず、実際の症例を基に作られたシナリオ（例えば、口唇口蓋裂で生まれた赤ちゃんの治療を、両親の苦悩と喜びに寄り添いながらどう行っていくか）を読ませ、そこから事実を抽出して問題を見だし、解決策（仮説）を考えさせます。この段階では、学生たちには必要な知識がないので、分からないことがたくさん出てきます。そこで、それを学習課題として設定させます。ここまですべてを1コマで行います。

次に、次週までに、それぞれが教室外で追加情報を集めます。そして、次の授業時間には、そのリサーチ結果を持ち寄って議論し、新しく得た知識と既存の知識を統合することによって、最初に立てた解決策について検証します。このようなプロセスで、問題発見・解決能力や自己学習能力の獲得と、学んだ知識の統合を目指すのがPBLです。

これは非常に良いやり方だと思いますが、評価は難しいとされてきました。私たちは、「改良版トリプルジャンプ」という方法を新たに開発しました。グループで学習しているのを教員が1対1の口頭試問のようなかたちで評価するトリプルジャンプという方法を修正したものです。これは三つのステップからなります。



まず教室の中で学習課題を設定し、次に図書館などで調べて、また教室で解決策（仮説）を検証するというステップⅠ、Ⅱは全てワークシートで行い、最後のステップⅢでは模擬患者（教員）を相手に解決策を実行させ、その後すぐにフィードバックを行います。つまり、ステップⅢは評価でもあり、指導でもあるわけです。学習成果はルーブリック得点の向上というかたちでも表れ、学生も「将来現場に出たときに役立つ」「調べたことが自分のものになった」など、十分な手応えを感じ取ってくれたようです。