

「学習成果」の設定と評価

—アカデミック・スキルの育成を手がかりに—

「学習成果」の設定と評価

—アカデミック・スキルの育成を手がかりに—

立教大学

大学教育開発・支援センター

はじめに

このシンポジウムは、当センターが学習成果の指標としての「ルーブリック」の開発という課題に直面したことから構想されました。「そもそもルーブリックとは何ぞや？」という初歩的な疑問からスタートして、何とか作成に取り掛かったものの遅々として進まず、「やはり専門家のお話と先行事例について詳しく伺いたい」という渴望がこの企画を生み出したのです。

京都大学高等教育研究開発推進センターの松下佳代先生に、ルーブリックを主題にしたご講演をお願いしたところ、ルーブリックの解説を含むパフォーマンス評価についてのお話ならお引き受けいただけるというご回答をいただきました。その時に、ルーブリック「だけ」しか考えていなかった己の不明を恥じました。学習成果とは何か、それをどう設定し、どう評価すべきか、その方法はいかにあるべきかといった総論に思いが至っていなかったのです。そのような総論的な理解がなければ、有効なルーブリックを作成し、活かしていくことは難しいでしょう。ご講演のお願いに際して、われわれをお叱りになることもなく、心優しく導いていただけたことは大変ありがたいことでした。

松下先生からは、この「学習成果とは何か、それをどう設定し、どう評価すべきか、その方法はいかにあるべきか」という問いに対して、本シンポジウムで明晰なご回答をいただきました。特に、学習成果の評価方法を4つに分類してご説明いただいたことで、パフォーマンス評価の位置づけとルーブリックの役割が明確になりました。また、パフォーマンス評価が、これまで大学の教員が経験則（主観的な評価）で行ってきた行為を、可視化しようとする手段であることもはっきり理解できました。

さらに、松下先生に新潟大学歯学部の小野和宏先生をご紹介いただき、事例報告をお願いすることができました。小野先生からは、まさにルーブリックを用いたパフォーマンス評価の実践にあたっての率直な「苦労話」をお伺いすることができました。たとえ付き合いの長い同僚や、他大学の友人との間であっても共有が難しい「苦労話」を伺うことができたのは、先生の教育への真摯な姿勢と寛大なお心があってのことです。具体的な苦労は「評価の信頼性」を高めるための試行錯誤にあったかと存じます。しかし、その試行錯誤こそが、大学教育の質を

高め、学生をさらに成長させていくことにつながるということがよく理解できました。

「大学で学習成果?」「なぜそこまでの必要があるのか?」「そもそも、そんなことが本当にできるのか?」「大学にはふさわしくないのではないか?」という声が聞こえてきそうなテーマです。一昔前なら、その通りなのかもしれません。しかし、様々な試行錯誤による教育の質の改善は、今まさに大学がやらなければならないことになったということだと思います。

ところで、ルーブリックという語は、キリスト教の祈祷書などへの書き込みに使われた赤いインク（ラテン語の rubrika）に由来しているといわれています。つまり、ルーブリックはキリスト教、そして立教大学と多少なりとも縁があったのです。本学において、ルーブリックに関連するシンポジウムを、お二人の寛大な先生方のご協力をもって開催できたことに心から感謝いたします。

大学教育開発・支援センター 副センター長、経済学部准教授

小澤 康裕

はじめに	
プログラム	9
趣旨説明 小澤 康裕 氏	11

第1部：講演

「学習成果」の設定と評価

—パフォーマンス評価を中心に—	19
	松下 佳代 氏

はじめに	19
1. 学習成果とは何か	19
2. 学習成果の設定と評価	21
3. 本質的学習成果 (ELO) と学士力	22
4. アカデミック・スキル	23
5. 学習成果の評価の構図	24
6. 学習成果の評価は、どうデザインし、どう実施すればよいのか	26
7. オルターナティブ・アセスメントの方法	27
8. パフォーマンス評価とルーブリック	29
9. ルーブリックとは	30
10. ルーブリック事例	32
1) AAC&Uの事例 —VALUEルーブリック—	
2) 関西国際大学の事例 —コモンルーブリック—	
3) 山口大学の事例 —コモンルーブリック—	
4) ユタ州立大学の事例 —歴史学科卒業論文用ルーブリック—	
5) 新潟大学歯学部事例	
【スライド①】	38

第2部：事例報告

「大学学習法」の学習評価実践

—レポート評価に焦点をあてて—	61
	小野 和宏 氏

はじめに	61
1. 学士課程をどうみるか	61
2. 「大学学習法」という科目	63
3. パフォーマンス評価への着目	63
4. ルーブリックの開発を始める	64
5. ルーブリックの改訂1 —観点を見直す—	65

6. 評価の信頼性の検討	66
7. ルーブリックの改訂2 ―信頼性分析を踏まえレベル数を増やす―	68
8. レベル数の設定と評価者間信頼性	68
9. 評価者によって異なる判断基準	69
10. 締めくくり	70
【スライド②】	73
【別紙資料】	80

第3部：講演

「学習成果」の設定と評価

―パフォーマンス評価を中心に― (2)	85
---------------------	----

松下 佳代 氏

1. ルーブリックの効用？	85
1) ぶれない？	
2) 時短？	
2. 規準の設定と基準合わせが重要	86
3. 〈誰が〉評価するのか	87
4. 〈学習としての評価〉という考え方	88
【スライド③】	90

質疑応答・ディスカッション	95
---------------	----

閉会挨拶 原田 久 氏	106
-------------	-----

「学習成果」の設定と評価

—アカデミック・スキルの育成を手がかりに—

2014年10月23日（木）開催
シンポジウムの記録

プログラム

「学習成果」の設定と評価

—アカデミック・スキルの育成を手がかりに—

趣旨説明 小澤 康裕 氏

(大学教育開発・支援センター 副センター長、経済学部准教授)

第1部：講演

「学習成果」の設定と評価

—パフォーマンス評価を中心に—

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)

第2部：事例報告

「大学学習法」の学習評価実践

—レポート評価に焦点をあてて—

小野 和宏 氏 (新潟大学 歯学部教授)

第3部：講演

「学習成果」の設定と評価

—パフォーマンス評価を中心に— (2)

松下 佳代 氏

質疑応答・ディスカッション

閉会挨拶 原田 久 氏

(大学教育開発・支援センター センター長、副総長、法学部教授)

司 会 御手洗 明佳 氏 (大学教育開発・支援センター 学術調査員)

日 時 2014年10月23日(木) 18:30 ~ 20:30

場 所 池袋キャンパス M301教室 (マキムホール3階)

「学習成果」の設定と評価

—アカデミック・スキルの育成を手がかりに—



司会

大学教育開発・支援センター 学術調査員

御手洗 明佳 氏

○**司会** では、お時間が過ぎましたので、始めさせていただきます。本日は、遅い時間で、さらに天気も芳しくない中、お集まりいただきましてありがとうございます。「『学習成果』の設定と評価—アカデミック・スキルの育成を手がかりに—」と題して、立教大学大学教育開発・支援センター主催のシンポジウムを始めさせていただきます。

私は、司会を務めさせていただきます、大学教育開発・支援センターの学術調査員、御手洗明佳と申します。よろしく申し上げます。

まず、趣旨説明として、当センターの副センター長、経済学部准教授の小澤康裕先生にお話ししたいと思えます。小澤先生、よろしく申し上げます。

■趣旨説明

○**小澤** 皆さん、こんばんは。本日はお忙しい中、お集まりいただきありがとうございます。私は、大学教育開発・支援センターで副センター長を務めております小澤です。どうぞよろしく願いいたします。

本シンポジウムのタイトルにあります「学習成果」については、平成24年8月に、「質的転換」答申というものが中教審から出されました。その中で、学習成果の把握という目標が掲げられており、それが大学教育の質保証につながる重要な課題であると認識されております。このような背景も踏まえまして、本シンポジウムでは、学習成果の設定と評価ということに焦点を当てております。京都大学の松下先生からは、パフォーマンス評価というものに注目しつつ、学習成果と

いう概念や、その評価のバックグラウンドについてご講演いただきます。続きまして、新潟大学の小野先生から、具体的な学習評価の実践についてお話をいただく予定です。

立教大学におきましては、これまで当センターで、「Master of Writing」、「Master of Presentation」といった学習支援教材を作成しております、アカデミック・スキルを育成するというのに力を入れてきました。現在は、こういったツールも使用しながら、レポートをどう評価するか、また、プレゼンテーションをどう評価するか、という課題を共有し、ルーブリックの作成を行っております。ルーブリックについては、お二人の先生のお話の中にも出てくると思いますので、私のほうからは詳しくは申し上げませんが、作成するのがなかなか難しいと感じております。どのような観点で学生のレポートを評価したらいいのかという点が定まらず、大変苦労しております。つくるならば、できるだけよいルーブリックをつくりたいと思ひまして、専門家でいらっしゃる松下先生と小野先生から参考になるお話を伺えたらなと考えています。

これから30分、まず松下先生からお話をいただきまして、その後続けて、小野先生にお話をいただいて、再度、松下先生からお話をいただくという予定でシンポジウムを進めさせていただきたいと思ひます。その後、質疑応答の時間を設



大学教育開発・支援センター 副センター長、経済学部准教授
小澤 康裕

けておりますので、ぜひ皆さまから、活発なご質問、ご議論をいただきたいと思っております。

それでは時間になっておりますので、先生方のご講演、ご報告をお願いしたいと思えます。どうぞよろしく申し上げます。

○司会 小澤先生、ありがとうございました。

続きまして、ご講演と事例報告に移る前に、松下先生と小野先生について、簡単なプロフィール紹介をさせていただきます。

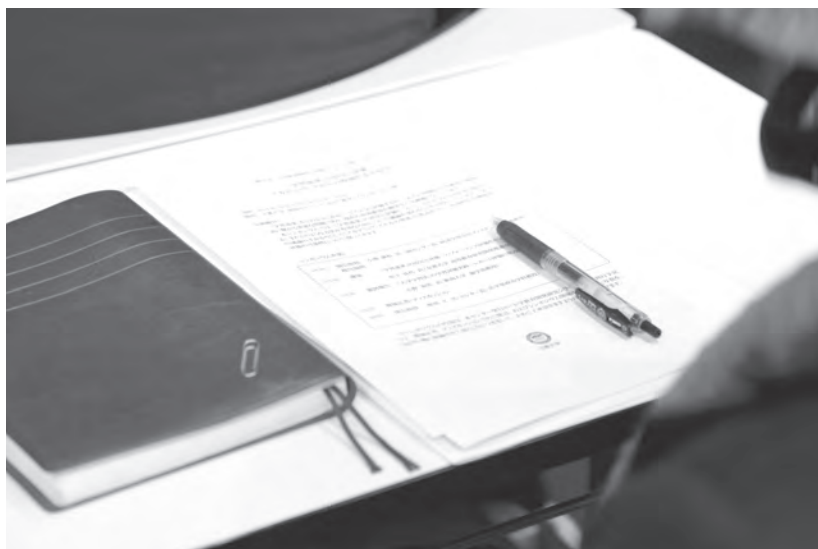
松下先生は、現在、京都大学の高等教育研究開発推進センターと京都大学教育学研究科の教授でいらっしゃいます。教育方法学と大学教育学がご専門で、多くの大学と交流しながら、精力的に共同研究をなさっています。また、諸外国でいわれるキー・スキルやコンピテンシー、我が国の初等・中等教育であれば「生きる力」、大学教育であれば「学士力」といった能力概念を、〈新しい能力〉として分析され、その評価の方法としてのパフォーマンス評価に注目された、ご著書や研究論文を執筆なさっています。現在、次期学習指導要領改訂に向けた文部科学省の資質・能力に関する検討会の委員でもいらっしゃいます。

事例報告をしていただきます小野和宏先生は、新潟大学歯学部の教授でいらっ



しゃいます。ご専門は口腔外科学と歯学教育学です。今回小野先生にさせていただくお話は、高次能力の評価方法の開発と研究についてです。小野先生は、ご自身が新潟大学の初年次教育科目である「大学学習法」を担当され、松下先生との共同研究として、アカデミック・スキルの育成と評価に取り組んでおられます。本日はそのご研究とご実践について、お話しいただきます。

それでは、松下先生お願いいたします。



第 1 部



[講演]

「学習成果」の設定と評価

—パフォーマンス評価を中心に—

京都大学 高等教育研究開発推進センター 教授

松下 佳代 氏

はじめに

ただいまご紹介にあずかりました京都大学の松下と申します。本日は、「学習成果」の設定と評価がテーマということで、御手洗さんからの依頼では、ルーブリックについて講演してくださいというお話だったのですが、後ほど言いますが、私自身は、ルーブリックというのはパフォーマンス評価という大きな評価の中の一部だと考えておりますので、副題は、「パフォーマンス評価を中心に」とさせていただきます。

1. 学習成果とは何か

まず、学習成果とは何かということについてですが、ご存じのとおり、学習成果というのは英語ではラーニング・アウトカムズのことです【スライド①-4】。これが注目されるようになってきたのは、日本では中央教育審議会の「学士課程教育の構築に向けて」という答申が2008年12月に出たあたりからなのですが、世界的に見れば、ラーニング・アウトカムズへの注目はそれよりかなり以前から世界共通の傾向になっていました。

先ほども出しましたが、この中教審の学士課程答申の中では、「学士力」というものが提案されたということが、大変関心を集めました。そこでは、「分野横断的に、我が国の学士課程教育が共通して目指す学習成果に着目したもの」として、この学士力が説明されています。

学習成果に注目するという事は、教員が何を教えたか—この「何を教えたか」というのは、往々にして、教えたつもりだけでも伝わっていないということもあったわけですが—、その何を教えたかということから、学生が実際に何を学んだかに目を向けるという、こういう転換が含まれていたわけです。

この「何を学んだか」ということですが、単に何を知り、理解したかということだけではなくて、実際に何ができるようになったのかという、「to know」だけではなく「to do」も含むような形で学習成果をとらえるようになったというのが、近年の大きな傾向です。

では、学習成果というのはどういう意味なのかということをご説明したいと思います【スライド①-5】。ラーニング・アウトカムズなのだから、ラーニングのアウトカムズ、学習の成果、あるいは結果なのだろうと普通には思うわけですが、単なる学習の結果ではないということを、まず強調しておきたいと思います。

最初に定義を見ておきますと、例えば OECD では、「学習者が学習プロセスの終了後に、何を知り、理解していて、また何をやってみせることができると期待されているかについての記述」として、「やってみせる」というところは「demonstrate」という英語が使われています。ここで、「…できると期待されているかについての記述」となっていることに、注目していただきたいと思います。学士課程答申の定義もほぼ同じような形で行われています。用語解説にあります。見比べていただければほぼ同じだということがおわかりになると思います。OECD はもっと以前から似たような定義をしていますので、学士力もこのような海外の動向を下敷きにしたのではないのでしょうか。

ここに示されているように、学習成果には2つの意味があるのですね【スライド①-6】。まずラーニング・アウトカムズということから私たちが考えてきたのは、結果としての、つまりラーニングした、学習した結果という意味ですね。その意味で使うときには、それを明確にするために、achieved learning outcomesという言葉が使われることもあります。これは結果としての学習成果であり、それは評価対象になるわけです。

近年加わった、もう1つの重要な意味は、目標としての意味です。この意味を明確にするときには、intended learning outcomesという表現が使われます。intendedの代わりに、先ほどの定義の中で「期待されている」という言葉があ

りましたが、expected とか、あるいは desired とか、そのような語がつけられることもあります。しかし、これらの語を全く使用せずラーニング・アウトカムズと言う言葉のみで、この意味をさすことも増えてきています。つまり学習成果というのは、単に結果ではなくて、目標を明確化するときにも用いられる概念であるということです。どのような学習成果の獲得を目指して教育、学習を行うかという、そういう目標という意味があるということですね。

2. 学習成果の設定と評価

では、このような意味での学習成果をどう設定し、どう評価すればよいのかということをお話していきたいと思います【スライド①-7】。学習成果の設定と評価、これは先ほど世界的に共通する傾向だと申し上げましたが、1つの例として、アメリカの全米カレッジ・大学協会（AAC&U）のものを参考に引用したいと思います。AAC&U というところは、全米の1,300 ぐらいのコミュニティ・カレッジからいわゆる研究大学までが含まれており、個人が参加するのではなく、機関として大学とかコミュニティ・カレッジが参加する、そのような組織です。特に、教養教育について議論したり、さまざまな提案を行ったりする団体です。AAC&U は、学習成果に関して、LEAP（リープ）プロジェクトという取組みを行っています。LEAP プロジェクトで議論されたのは、教養教育における essential なラーニング・アウトカムズです。これは、「中枢的な」とか、「本質的な」とか、「不可欠な」とか、あまりいい訳が思い浮かばなかったのですが、そのようなラーニング・アウトカムズとして、どのようなものがあるのかということを議論する、それがこの LEAP プロジェクトの目指すところでした。

ですので、アメリカの liberal education、すなわち、日本の1年生、あるいは2年生ぐらいまでの教養教育ということではなくて、ほぼ4年間、学士課程教育全体で身につけるべき学習成果—ここでは言うまでもなく、intended、expected なラーニング・アウトカムズですが—、そのようなものとして、何が挙げられるのかということ議論して、提案したのが、この LEAP プロジェクトでした。次のスライドでお見せしますが、学士力と共通点が多くあります。

一方、それをどのように評価していけばいいのかを明らかにすることを目指して行われたのが、この VALUE プロジェクトです。VALUE は、もちろん「価値」

とひっかけていますが、Valid Assessment ということ、妥当性のある評価を提案するという事です。この中で提案されたものが、本質的な学習成果を評価するためのルーブリックで、「VALUE ルーブリック」と呼ばれています。

これは、例えばリテンション率とか、卒業率とか、そういった間接指標だけで学習成果を評価するのでは不十分だということ、それからまた、2000年代に入ってから、アメリカでは標準テストで学習成果を測るということが非常に広がっていったのですけれども、その標準テストへの批判という意味も込めて、このVALUEプロジェクトは行われました。今日はどういう批判が行われたかということまでは立ち入ることができませんが、そういうものとは異なる形で学習成果を評価するというのが目指されたわけです。

13. 本質的学習成果 (ELO) と学士力

では、まず本質的な学習成果 (Essential Learning Outcomes) ということ、どういうものが抽出されたのかということですが、表の左側に示しました【スライド①-8】。そして右側が2008年の学士課程答申で出てきた学士力です。こうやって見比べていただくと、非常に似通っていることがお分かりいただけだと思います。一番上のカテゴリーが各学問分野、専門分野で身につける知識とか理解です。それから、その下のところは、そういう各専門分野に特殊なものではなくて、もっと幅広い—ここでは汎用的という言葉が使われていますが、分野横断的と言ったほうが正確かもしれませんが—、そういったものが並んでいます。本質的学習成果では、「知的・実務的スキル」という表現がされていますが、学士力のほうでは「汎用的技能」という言い方がされていますね。

それからこの「個人的・社会的責任」というところは、学士力では、「態度・志向性」と読み替えられています。それから、「統合的学習」ということですが、アメリカの教養教育では、4年生でキャプストーン科目—日本で言うと、卒業研究とか卒業論文に当たるような科目—も含まれていますので、そのような学士課程で学んできたいろいろな知識・技能などを統合して行う学習のことです。学士力でも、ほぼ似通った表現がされています。

よく見ると、多少違うところがあるのですが、ほぼ同じということで、おそらく学士力は、この本質的学習成果というものをある程度下敷きにしなが

案されたのだらうと思います。後ほど AAC&U の VALUE ルーブリックについて詳しくお話ししますが、今ご紹介した LEAP プロジェクトや VALUE プロジェクトは、日本の高等教育政策にも影響を与えているということを念頭に置いていただければと思います。

【4. アカデミック・スキル

今回のシンポジウムの副題である、アカデミック・スキルについてですが、アカデミック・スキルについては、今日お見えの方は、皆さんよくご存じかと思えます【スライド①-9】。一応、定義らしきものをおきたいと思えます。大学で、そして、大学での学びだけではなく、これから人生でさまざまな学問的・知的作業を行うために必要なスキルととらえられると考えています。

より具体的に言えば、自ら考え、調べ、論じ、書くためのスキルです。本質的学習成果のほうで言えば知的・実際的スキル、学士力のほうで言えば汎用的技能、そういうものの一部と言えるのではないかと思います。それは基礎的でもあり、また分野横断的でもあると考えられます。

私は、こういうふうな模式図を考えているのですが【スライド①-9】、例えば問題解決ということがよく言われます。その問題解決能力というのは、分野



によっていろいろな形を取るはずなのです。後ほど、新潟大学の歯学部のお話がありますが、例えば新潟大学の歯学部のPBLでやっているような問題解決と、数学での問題解決は全く違う。けれども、共通しているところがあるはずなのです。ですので、各分野で全く同じではないけれども、模式図の中心に示したような、共通する部分があると。それは多分、すごく基礎的な部分だと思うのです。それを初年次のところで特に取り出して教えるのがアカデミック・スキルではないかな、と私は考えています。

15. 学習成果の評価の構図

ここまで、学習成果についてお話してきました。では、このような学習成果を、どのように評価すればいいのかについてですが、私は現在行われている学習成果の評価というものを、大きく2つの軸でとらえることができていると思っています【スライド①-11】。1つは直接評価と間接評価という軸です。もう1つは、心理測定学的パラダイムと、それに代わる、または、それを補うという形で示されるオルターナティブ・アセスメントのパラダイムです。

中教審の質的転換答申の中で、これまでのアドミッション、カリキュラム、ディプロマという、3つのポリシーに加えて、もう1つアセスメント・ポリシーというものが提案されています。その中で、特に学習成果の評価にあたっては、学修行動調査とか、アセスメント・テストとか、ルーブリック、学修ポートフォリオなどを用いるということが書かれています。このような評価方法の違いというのも、この2つの軸を置いて考えると、その違いが浮き彫りになるのではないのでしょうか。それを図として表したものがこちらです【スライド①-12】。間接評価と直接評価、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4つの象限ができます。

例えば、この第1象限のところには、ミニッツペーパーを挙げました。ほかのところだと、例えば、先ほど出た学修行動調査というのは、この第2象限のところの学生調査の一部になります。それからアセスメント・テストは、標準テストや客観テストの形で行われる第3象限のところ当てはまります。そして、ルーブリックや学修ポートフォリオは、この第4象限のところに入ると考えています。

それぞれの評価方法についてももう少し詳しく見ていきましょう。ミニッツペー

パーはお使いになっている方もいらっしゃるかもしれませんが【スライド①-13】。もともと英語では One Minute Paper ですので、ミニットペーパーですね。例えば、「今日の授業で学んだことの中で一番面白かったことはこういうことだった」という事柄を自由記述の形で、毎時間の授業の後に書かせるというタイプのもですね。タイプⅠと言っていいと思います。

次に、これは同志社大学の山田礼子先生が現在実施している新入生学習調査の中の1つの項目なのですが、[あなたは現在、以下の能力や知識はどの程度身につけていますか]ということで、例えば一般的な教養とか、分析や問題解決能力とか、そういったものを「とても少ない」から「とても多い」までの5段階で、学生自身に評価させて答えさせるというものです【スライド①-14】。これがなぜタイプⅡなのかは、また後ほどお話ししますが、先ほどのミニットペーパーとは全然違う形で学習成果を見ようとしているのはお分かりいただけると思います。

それから、もう1つ、これは皆さんおなじみの大学入試センター試験ですが、こういう形で、非常に細かく評価項目を分割して、機械でもできるくらいの客観性を持たせて評価をするという形のもです。これはタイプⅢになると思います【スライド①-15】。

そして、大学でよく使われるレポートとか卒業論文、プレゼンテーション、口頭試問、あるいはいろいろな形で行われるプロジェクト、美術系での卒業制作、あるいは教員養成系での模擬授業だとか、法学部、法科大学院での模擬裁判など、大学というところはいろいろな形で学習成果の評価を行っているわけです【スライド①-16】。私は、こういうものはタイプⅣに入っていると思っています。

大学というところは、ペーパーテスト以外の多様な評価の宝庫だと思うのですが、このような形での学習成果の評価をどうデザインし、どう実践すればよいのかということについては、これまでの経験則でやってきている部分が多いように思われます。ここにメスを入れようとするのがパフォーマンス評価であり、その一部としてのルーブリックだと考えています。

6. 学習成果の評価は、 どうデザインし、どう実施すればよいのか

まず、この間接評価と直接評価の違いということですが、間接評価というのは、学生の学習行動や、学習成果についての自己認識によって行う評価です。学生自身に、自分はどのくらいの能力を身につけているかを自己報告させることを通じて、学習成果を間接的に評価しているものです【スライド①-17】。学生の自己報告によるという点で、ミニツツペーパーも、学生調査も間接評価であるということがお分かりいただけたと思います。それから直接評価というのは、先ほど demonstrate という言葉がありましたけど、学生の知識や技能を何らかの形で表に出させて評価する。レポートとか、卒業論文とか、プレゼンとかもそうだと思うんですね。これが直接評価です。間接評価は、「何ができると学生が思っているのか」、直接評価は、「実際に何ができるのか」ということで評価をするというものです。

もう1つの評価の2つのパラダイム【スライド①-18】ですが、これについては、詳しく述べるとこれだけで随分時間がかかってしまうので、ごく簡単にお話ししたいと思います。このように、ちょっと荒っぽいのですが、分かりやすく整理してみました。心理測定学というのは、今でも評価論のメインストリームで



す。特にアカウンタビリティを目的にしたような評価とか、あるいは入試のように非常に厳密性が求められるような場合には、今でもメインで使われているものです。先ほど見たように、学生調査にしても、それからセンター試験にしても、細かく項目を分割して評価をしていきます。そこでの評価データは、量的なデータであり、統計的な分析にかけられます。評価基準は非常に客観的なものです。

それに対して、先ほどの、例えばミニッツペーパーもそうですし、それからレポートとか、模擬授業とか、プロジェクトとかもそうなのですが、評価項目が複合的なものであるということが分かっていたかと思うのです。特にタイプⅣのほうに当てはまると思うのですが、さまざまな知識とか技能を使いこなすことが必要な課題を、何か文脈を設定して行わせる。また、評価基準についていえば、普通、卒論などは3人くらいの教員で評価すると思うのですが、これは1人だけだと恣意的になるかもしれないので、主観的ではあるのだけれども恣意的にならないように、主観を突き合せるという意味での間主観性というのが、評価基準のところで見られる特徴です。評価データは質的なデータです。ここでの評価主体は実践者自身、つまり、教員、それから学生である場合もあります。

17. オルターナティブ・アセスメントの方法

最近提唱されてきている評価方法に、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があります【スライド①-19】。これらはすべてタイプⅣの評価に当たります。

「ポートフォリオ評価」については、多分ここにいらしている方は、ほとんどがご存じかと思います。ポートフォリオ、つまり紙ばさみに、その学習者の成長の記録となる資料を入れていって、そして一定期間経ったときに振り返って、その学習者の成長のプロセスを評価する、そういう方法ですね。これが学生ではなくて教員だったら、ティーチング・ポートフォリオになるわけです。それから「真正の評価」というのは、課題とか、それからその課題をやる文脈などが、非常に本物らしい、オーセンティックであるということを強調した評価です。さらに、「パフォーマンス評価」、これは実際に遂行させることを通じて行う評価です。先ほど、単に知り、理解しているだけではなくて、実際に「to do」、することが求められるようになってきたと言いましたけれども、実際に何かをやってみるとか、あ

るいは何かをつくらせるとかレポートなどもそうだと思うのですが、そういうものを通じて直接的に評価する。それがパフォーマンス評価なのですね。ですので、真正の評価とパフォーマンス評価は、かなり重なっているところがあるのです。強調点の違いと言ってもいいかなと思います。

3つの評価方法の関係をどういうふうに見るかというのは、論者によっていろいろ違いがあるのですが、私はこんな感じでまとめています【スライド①-20】。評価の中にいろいろなやり方があるのですが、パフォーマンス評価の場合だったら、何らかのパフォーマンスをさせる。その中で特に、本物らしい課題とか文脈の中で行われるのが真正の評価だと。例えば自動車の仮免と路上試験を考えていただくと思ふのですが、どちらも実際に運転させてみるのですが、自動車学校の中のような、非常に限られた場面でやるのか、それとも、本当に路上の、いろいろな条件が統制されていない中でやるのか。どちらもパフォーマンス評価なのだけれども、路上試験のほうがより真正の評価とすることができます。それ以外にも、私たちは客観テストもやりますし、ワークシート、あるいは学生調査で自分がどのように答えたかなども、学習の証拠、エビデンスになります。そういうものを束ねたものがポートフォリオですね。それを一定期間経ったところで振り返ってみて、自分の成長を評価するのがポートフォリオ評価ということになります。



ルーブリックというのは、今お話ししたようなオルタナティブ・アセスメントをやるときに使われる評価基準の1つです【スライド①-21】。最近、「ルーブリック評価」という言葉も、特に大学教育では使われるのですが、私はルーブリックだけひとり歩きさせてはだめで、あくまでも今申し上げたように、パフォーマンス評価とか真正の評価とか、そういった評価全体のシステムの中でルーブリックを考えることが必要だと考えています。

18. パフォーマンス評価とルーブリック

今、パフォーマンス評価の中でのルーブリックと言いましたけれども、それはどのような関係にあるのかということをごここからお話ししていきたいと思えます【スライド①-23】。パフォーマンス評価についての問い、誰が、何を、何によって、何をエビデンスとして、どのように評価するのか。これは多分、どんな評価でも立てられる問いだと思うのですが、パフォーマンス評価では、教員や学生が、学習成果を、実演とか作品といったようなパフォーマンスによって、ルーブリックを使って評価をするということになります。評価基準としてはルーブリック以外に、チェックリストが使われる場合もあります。

今申し上げたことを概念図でお示しします【スライド①-24】。パフォーマンスの対概念はコンピテンスです。チョムスキーの言語学理論をご存じの方は、彼が、linguistic competence と linguistic performance、すなわち言語能力と言語運用という対概念で理論を構築しているのをご存じかと思えます。どんな能力も、それ自体を観察することはできません。例えば各地に力石^{ちからいし}というものが残っているようなのですが、力石というのは、若者にその石を持ち上げさせて、一人前になったかどうかを試すために使われたそうです。一人前の男と言えるだけの力があるかどうか、それはそのままでは見ることができないので、その力石を持ち上げさせるというパフォーマンス、実演をさせることによって、観察可能にして、ちゃんと持ち上げられたら、一人前の男になったと、そういうふうに解釈するということなのですね。このように、観察不可能な能力を、いったんパフォーマンスという形で観察可能にして、そしてそれを解釈して、能力が身についているかどうかを評価していく、これがパフォーマンス評価の概念図です。

その可視化をするときに、パフォーマンス課題を使い、そして解釈・評価する

ときにルーブリックを使うということになります。パフォーマンス課題は、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課題で、先ほど言いましたように複合的なものです【スライド①-25】。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です。これは評価課題の分類【スライド①-26】で、パフォーマンス評価がどこに位置づくのかを示しています。

私はもともと初等・中等教育の研究をやっていたので、それと比較して強く感じることもなのですが、先ほどお話ししたように、大学教育というのはたくさんパフォーマンス課題があるな、本当に宝庫だなと思うのですね。ですが、評価基準については、ほとんど主観に委ねられていました。ただ、主観に委ねられていたといっても、大学教員の場合は、ほとんどの人がその分野のエキスパート、専門家なわけですね。ですから専門家の鑑識眼というのを持って評価していたので、大きく問題になることはなかったと思うのです。けれども、その先生にしか分からない形で、主観に基づいて評価をするというだけでは、ほかの人に対して説明をしたり、あるいは教員同士で共有したりすることが難しい。ですので、これまで自分の頭の中でやっていたことを、表にあらわすということがルーブリックの意味としてあると思います。

【9. ルーブリックとは

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」のことで、このようなマトリックスの形を取ります【スライド①-27】。段階的というのは、ふつう4段階から6段階ぐらいのレベルを設定して、パフォーマンスを段階的に評価するということです。多面的というのは、1つのパフォーマンスをいろいろな点から見るということです。日本語では「観点」と言う場合が多いですが、英語では criteria、あるいは dimension と言ったりします。ルーブリックの特徴というのは、先ほどの学生調査の5段階評価みたいに単に1から5までで評定してくださいというのではなくて、実際に、例えば観点4のレベル1だったら、どういうふうな特徴が見られたら、観点4はレベル1なのかということを言葉で表すということです。これが記述語、descriptor です。先ほど申し上げたように、ルーブリックというのはあくまでも専門家の鑑識眼をイメージ化し、共有できるようにするためのツールだということを強調し

ておきたいと思います。

ただ、これがまたルーブリックの特徴なのですけれども、いったんこのようなルーブリックを使うと、パフォーマンスの質は量としてあらわすことができるようになります。これは、一応いい面だと言っていていいかなと思いますけれども、そうすると、もともとは質的データだったものが、量的データとして扱われるようになります。ルーブリックが今、いろいろなところで使われるようになった。しかもオルタナティブ・アセスメントだけでなく、心理測定学的パラダイムのほうでもルーブリックが使われるようになってきているのは、質を量に変換するという働きをルーブリックが持っているからだと思います。

ルーブリックのタイプですが【スライド①-28】、まず、ルーブリックの構造に注目すると、今ご紹介したのは幾つかの観点を設定した分析的ルーブリックでしたけれども、観点を設定しないで、ホリスティックに評価をする全体的ルーブリック、ホリスティック・ルーブリックというものもあります。

それからスコープですね。これはある領域で、一般的に適用できるものなのか、特定の課題だけに適用できるものなのかということです。今日のお話の範囲では、この違いはさほど話題になりませんが、実は新潟大学でやっている PBL、Problem Based Learning の評価では、この二つを使い分けています。

今日お話しする中では、ルーブリックのスパン、対象期間というものが重要で



す。単に、あるプロジェクトや、あるレポートを評価する、採点するためのルーブリックだけではなくて、複数年にまたがって使われる長期的なルーブリックも、今、結構使われるようになってきています。こういったルーブリックの事例について、この後お話ししていきたいと思います。

10. ルーブリック事例

ご紹介するルーブリック、まず AAC&U の VALUE ルーブリックです【スライド①-30】。それから、日本でかなり早いうちからルーブリックを取り入れて、実践をさせてきている関西国際大学のコモンルーブリックですね。それから山口大学。これも今、私に関わらせていただいているのですが、初年次科目「山口と世界」ーローカルとグローバルを結びつける、そういう初年次科目を全学でやられているのですけれどもーその中で開発されているルーブリック。それからユタ州立大学の歴史学科で使われている卒業論文用ルーブリック。なぜユタなのかというのは後ほどお話ししたいと思います。それから新潟大学でのルーブリックですね。これは後ほど、小野先生からご紹介があります。

これらは、それぞれちょっとずつ階層が違うのです。AAC&U の場合は、先ほど言いましたように 1,300 もの大学が共有できるものとして、この VALUE



ルーブリックは提案されました。しかも科目横断的に使われるルーブリックです。関西国際大学の場合は、全学で科目横断的に共有するものとして作成されています。それから山口大学やユタ州立大学、新潟大学の歯学部の場合は、それぞれ特定の科目において教員間で共有できるものとして提案されています。

1) AAC&U の事例—VALUE ルーブリック—

まず AAC&U の VALUE ルーブリックですが、先ほどの本質的学習成果がほぼ対応していて、もともとは 15 領域で提案されていました【スライド①-31】。その中には、批判的思考やチームワークなどが含まれています。その後、もう 1 つ、グローバル学習のルーブリックが加わって、今は 16 領域になっています。

VALUE ルーブリックは学士課程 4 年間の長期的なルーブリックです。また、メタルーブリックというか、普通のルーブリックより一段、抽象度が高いルーブリックと言えると思います。というのも、ルーブリックはアメリカの大学でもう 20 年近く使われていると思いますが、もともといろいろな大学で使われていました。その大学で使っている人たちが集まって、共同開発のチームをつくって、さまざまな VALUE ルーブリックを作成したのです。ただ、先ほど申しましたように、2 年制のコミュニティ・カレッジから、例えばジョージタウンのような世界のトップ 10 に入るような研究大学まで、多様な大学が加盟している協会です。その全機関が同じルーブリックで評価できるわけがないのです。なので、VALUE ルーブリックをもとにしながら、各大学、学科、科目で、それぞれの文脈に合わせて、ローカライズ、あるいはモディフィケーションという言葉を使っていますけれども、そのような作業をすすめ、自分たちのルーブリックをつくる。そのもとになるのが、この VALUE ルーブリックなのです。そのようにして、共通性と多様性の両立を図ろうとしたわけですね。

これは文章コミュニケーションの VALUE ルーブリックです【スライド①-32】。私どものセンターの紀要に掲載された論文の中で 15 の VALUE ルーブリックの訳を掲載しています。ご興味がおありの方はそちらをご覧くださいと思います（注）。ダウンロードもできます。ここでは途中は省いていますが、ベンチマー

(注) 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。

グは1年生として入学したときの水準、キャップストーンは4年生で卒業するまでに達してほしい水準を示しています。両者の間の水準はマイルストーンという言葉であらわされています。1に満たなければゼロなので、これは5つの観点と5段階のレベルからなるルーブリックということになります。4年間で作成するいろいろな文章コミュニケーションについての評価基準を、全部これにまとめているので、かなり抽象度が高くなっています。ですから、このままではレポート評価に使うというわけにはいかないだろうと思います。

実際、去年、このVALUE プロジェクトの責任者だった Rhodes 先生が来日されたときに紹介されたのは、このようなルーブリックでした【スライド①-33】。ある大学の1年生対象のバイオロジーの科目のレポート評価用ルーブリックとして、こういうものを紹介されたのです。これを見ると、各観点というのは、VALUE ルーブリックのいろいろなところから持ってきてつくられていることが分かります。そして1年生のレポートなので、3段階でいいと考えられたのでしょう。モディフィケーション、ローカライズして使われています。いろいろなVALUE ルーブリックから必要だと思われる観点を引っ張ってきて、レベルの数や表現も変えたりして使われているわけです。

VALUE ルーブリックのまま使っている大学もごくまれにはありますが、ほとんどの大学はこのようにアレンジして使っています。



2) 関西国際大学の事例ーコモングループリックー

次に、関西国際大学のコモングループリックを見てみましょう【スライド①-34、35】。関西国際大学の場合は VALUE ループリックを参考にして、全学共通のコモングループリックをつくられています。科目の違いを超えて、どの科目でも身につけさせようとしているスキルとして、ライティングとプレゼンテーションとリサーチ、この3つのスキルを選んでいきます。そしてここが特徴なのですが、全体は0～5までの6段階で構成されています。そのうち0～3の部分を下位学年版、2～5の部分を上位学年版としています。下位学年だったら、3までいけば、ほぼちゃんとできたと言える。上位学年は5までいってほしいと。こういうふうには、1つのループリックで4年間全体をカバーするのではなくて、上位学年版と下位学年版に分けて使われています。

この前、インタビューに伺ったのですが、試行を重ねて、2～3年ぐらいかけて何回もつくり直して、ようやく今、大体使えるものになってきたかなということ、学長の濱名先生はおっしゃっていました。全授業の大体6割くらいで使われているということでした。

3) 山口大学の事例ーコモングループリックー

次に山口大学のコモングループリックです。昨年度、全学必修の初年次科目に「山口と世界」という科目が設けられました【スライド①-36】。ただ、全学必修で、「山口と世界」という科目名は同じ、高次の到達目標も同じ、という条件のなか、それを担当する先生方のご専門はばらばらという授業です。つまり、先生方が、自分の専門性を生かしながら、「山口と世界」というテーマに迫るという科目なのです。ですので、先ほどの共通性と多様性の両立と同じ話になるのですが、共有できる部分をまずコモングループリックとしてつくり、それを各先生が、自分の授業に合わせてローカライズ、カスタマイズして使うということになっています。同じコモングループリックといっても、先ほどの関西国際大学の場合とは少し違うわけですね。学生が取り組む課題も、調査、研究、商品開発など多様なものが含まれているそうです。

4) ユタ州立大学の事例ー歴史学科卒業論文用ループリックー

ユタ州立大学の歴史学科の卒業論文用ループリックです。なぜユタかといいま

すと、ユタ州内に全部で6大学あるらしいのですが、ユタ州立大学歴史学科の取組みというのは、ユタ州立大学以外も含めた6大学の歴史学分野で共有するようになったのです【スライド①-37】。そして、その取組みが全米のアメリカ歴史学会にも注目されて、今では全米の歴史学の分野で、全く同じものではないですが、ルーブリックが提案されています。アメリカ歴史学会のホームページに入ると、それを見ることができます。

例えば、ユタ州立大学の場合、歴史学科全体で学生が身につけるべき学習成果を歴史学の知識、歴史学の思考、歴史学のスキルとして、大きく3つのカテゴリーを設定しています。それにさらに細かく規準を設定して、最終的な卒業論文の成績評価にも使えるようなルーブリックをつくっています。この8月に、国立教育政策研究所が、このユタ州立大学の取組みにおいて中心的な役割を果たしてこられた Daniel McInerney 先生を招聘されたので、京大でもシンポジウムにお招きしてお話を伺いました。まだ公開されていないもののようなので、見ていただくだけにしますが、英語で配られたものを私の方で訳してみました。このように、それぞれの規準ごとに重みづけされて、各レベルの点数が示されています。歴史学の知識は1つの規準だけですが、歴史学の思考や歴史学のスキルは、さらに細かくいくつかの規準に分かれています。歴史学なので、歴史学らしく、例えば「1次資料が使えているか」、「1次資料と2次資料をちゃんと使い分けているか」といった内容が入っていますね。学問分野の特徴をよく踏まえた上で、ライティングを評価しようとしていることがお分かりいただけるかと思います。この後お話しする新潟大学の場合は、ルーブリックの中に知識はあえて入れていませんが、それは初年次科目だからという面もあります。


5) 新潟大学歯学部事例

最後に、新潟大学歯学部のパフォーマンス評価の取組みについてお話しします。今開発しているのは、1つは「大学学習法」という初年次科目のレポート評価用ルーブリックです【スライド①-38】。もう1つは、2年生から5年生までの専門科目におけるPBL テュートリアルのためのパフォーマンス評価の開発です。こちらのほうはパフォーマンス課題の開発から関わっています。

今日はこの後、小野先生に、「大学学習法」のルーブリック開発について苦勞話をお話しいただくことになっております。ではここでいったん小野先生にバト

ンタッチしたいと思います。ご清聴ありがとうございました。

【スライド①-1】




立教大学教育開発・支援センター主催 秋期シンポジウム

2014.10.23

**「学習成果」の設定と評価
ーパフォーマンス評価を中心にー**

松下 佳代
京都大学・高等教育研究開発推進センター
matsushita.kayo.7r@kyoto-u.ac.jp

【スライド①-2】



OUTLINE

- 学習成果とは何か
- 学習成果の評価の構図
- パフォーマンス評価とルーブリック
- ルーブリックの事例
- 学習としての評価

2

【スライド①-3】



【スライド①-4】

学習成果への注目

- 学習成果 (learning outcomes) への注目
 - 世界共通の傾向
 - 日本では、中教審「学士課程答申」(2008.12)頃から「学士力」…「分野横断的に、我が国の学士課程教育が共通して目指す学習成果に着目したもの」
 - 教員が何を教えた(つもり)か
 - 学生が何を学んだか
 - 何を知り理解したか (to know)
+
 - 何ができるようになったか (to do)

4

【スライド①- 5】

「学習成果」の意味

* 単なる「学習の結果」ではない

- 定義
 - OECDによる定義
 - 「学習者が、学習プロセスの終了後に、何を知り、理解していて、また何をやってみせる(demonstrate)ことができると期待されているかについての記述」(OECD, 2009, p.3)
 - 学士課程答申での定義
 - 「プログラムやコースなど、一定の学習期間終了時に、学習者が知り、理解し、行い、実演できることを期待される内容を言明したもの」(「用語解説」)

5

【スライド①- 6】

つまり、学習成果には2つの意味がある

- 目標…intended/expected/desired learning outcomes
- 結果(評価対象)…achleved learning outcomes

では、学習成果をどう設定し、どう評価するのか？

6

【スライド①- 7】

学習成果の設定と評価—AAC&Uの場合—

*Association of American Colleges and Universities(全米カレッジ・大学協会)

- LEAPプロジェクト
 - *Liberal Education and America's Promise
 - 目標: 教養教育の本質的学習成果(Essential Learning Outcomes)
 - 学士課程段階で身につけるべき学習成果(意図された学習成果)
 - 「学士力」の元ネタ

- VALUEプロジェクト
 - *Valid Assessment of Learning in Undergraduate Education
 - 評価: 本質的学習成果を評価するためのルーブリック(VALUEルーブリック)を開発
 - 間接指標では不十分、標準テストへの批判

7

【スライド①- 8】

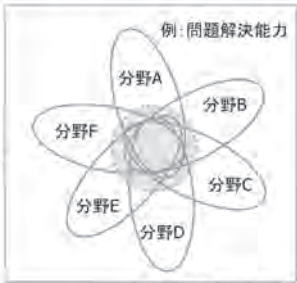
本質的学習成果(ELO)	学士力
人類の文化や自然界についての知識 ・科学、数学、社会科学、人文学、歴史、言語、芸術などの学習を通じて	知識・理解 ・多文化・異文化に関する知識の理解 ・人類の文化、社会と自然に関する知識の理解
知的・実際のスキル ・探究と分析 ・批判的思考・創造的思考 ・文章コミュニケーション・口頭コミュニケーション ・量的リテラシー ・情報リテラシー ・チームワークと問題解決	汎用的技能 ・コミュニケーション・スキル ・数量的スキル ・情報リテラシー ・論理的思考力 ・問題解決力
個人的・社会的責任 ・市民としての知識と関与(ローカル、グローバルに) ・異文化の知識と能力 ・倫理的な推論と行為 ・生涯学習のための基礎とスキル	態度・志向性 ・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
統合的学習 ・一般教育・専門教育での統合とより高度な達成	統合的な学習経験と創造的思考力 (AAC&U, 2007)

8

【スライド①- 9】

● **アカデミック・スキル** (cf. 佐藤他, 2006)

- 大学で、そして、これから人生でさまざまな学問的・知的作業を行うためのスキル
- より具体的には、「自ら考え、調べ、論じる」ためのスキル
- 「知的・実務的スキル」や「汎用的技能」の一部
- 基礎的であり分野横断的でもある




例: 問題解決能力

(松下, 2014a) 9

【スライド①- 10】

学習成果の評価の構図



10

【スライド①-11】

学習成果の評価のタイプ

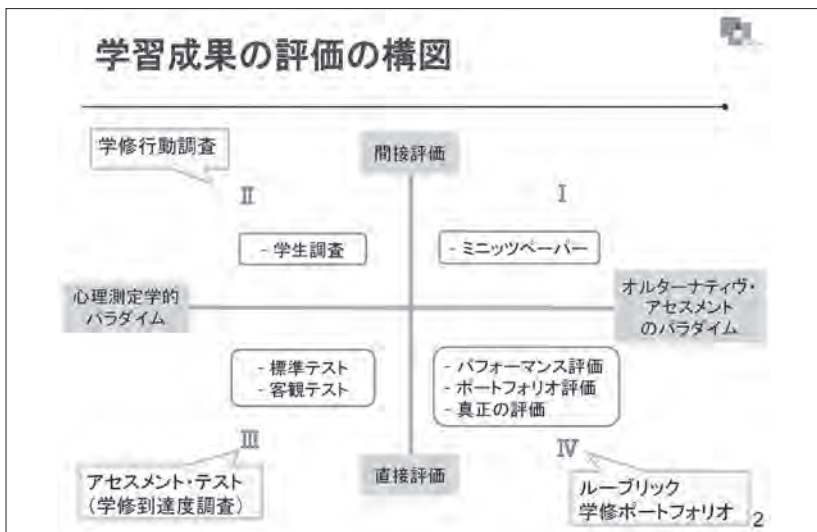
- 現在行われている学習成果の評価
 - 直接評価 vs. 間接評価
 - 心理測定的パラダイム vs. オルターナティブ・アセスメントのパラダイム

という2つの軸で捉えることができる (松下, 2012)

cf. アセスメント・ポリシー(中教審「質的転換答申」2012.8)
 「特に、成果の評価に当たっては、学修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、どのような具体的な測定手法を用いたかを併せて明確にする。」

11

【スライド①-12】



【スライド①-13】

【例】 ミニッツペーパー タイプ I

One-Minute Paper

- The most important/interesting thing I learned in class today was
- As a result of today's class discussion, I am confused about
- After doing the readings for class today, I have questions about
- I would like to spend more time discussing

Not necessary to respond to every statement

<http://www.nln.org/ce/mcgovern/sld047.htm>

13

【スライド①-14】

【例】 「新入生学習調査」 (山田他) タイプ II

(13) あなたは、現在、以下の能力や知識はどの程度身についていますか。
(それぞれ1つを選んで○をつけてください)

	とても多い	多い	それほど多い	少ない	とても少ない
1. 一般的な教養	++5..	++4..	++3..	++2..	++1..
2. 分析や問題解決能力	++5..	++4..	++3..	++2..	++1..
3. 専門分野や学問の知識	++5..	++4..	++3..	++2..	++1..
4. 批判的に考える能力	++5..	++4..	++3..	++2..	++1..
5. 異文化の人々に関する知識	++5..	++4..	++3..	++2..	++1..
6. リーダーシップの能力	++5..	++4..	++3..	++2..	++1..
7. 人間関係を構築する能力	++5..	++4..	++3..	++2..	++1..
8. 他の人と協力して仕事を遂行する能力	++5..	++4..	++3..	++2..	++1..
9. 異文化の人々と協力する能力	++5..	++4..	++3..	++2..	++1..
10. 地域社会が直面する問題の理解	++5..	++4..	++3..	++2..	++1..
11. 国民が直面する問題の理解	++5..	++4..	++3..	++2..	++1..
12. 文章表現の能力	++5..	++4..	++3..	++2..	++1..
13. プレゼンテーションの能力	++5..	++4..	++3..	++2..	++1..
14. 整理的な能力	++5..	++4..	++3..	++2..	++1..
15. コンピュータの操作能力	++5..	++4..	++3..	++2..	++1..
16. 卒業後に就職するための準備の程度	++5..	++4..	++3..	++2..	++1..
17. コミュニケーションの能力	++5..	++4..	++3..	++2..	++1..
18. 時間を効率的に利用する能力	++5..	++4..	++3..	++2..	++1..
19. グローバルな問題の理解	++5..	++4..	++3..	++2..	++1..
20. 外国語の能力	++5..	++4..	++3..	++2..	++1..

14

【スライド①-15】

【例】 センター試験 タイプⅢ

英 語

(解答番号 1 ~ 35)

第1問 次の問1・問2に答えよ。(配点 6)

A 次の問1・問2において、下線部(a)・(b)の単語のアセント(強勢)の位置が正しい組合せを、下の①~④のうちからそれぞれ一つずつ選べ。

問1 1

The smooth (a) surface of the (b) desert makes it possible for camels to transport goods easily.

① (a) surface (b) desert ② (a) surface (b) desert
 ③ (a) surface (b) desert ④ (a) surface (b) desert

問2 2

The coach (a) forget to tell the swimmer that his (b) record had been broken by a rival.

① (a) forget (b) record ② (a) forget (b) record
 ③ (a) forget (b) record ④ (a) forget (b) record

15

【スライド①-16】

タイプⅣ

【例】 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…

↓

このような形での学習成果の評価は
どうデザインし、どう実施すればよいのか？

16

【スライド①-17】

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？

17

【スライド①-18】

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、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

18

【スライド①- 19】

オルターナティブ・アセスメントの評価方法

- **ポートフォリオ評価** (portfolio assessment)
 - 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** (authentic assessment)
 - 学習者が、知識や技能などを使って、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
- **パフォーマンス評価** (performance assessment)
 - 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

19

【スライド①- 20】

● 3つの評価方法の関係

ポートフォリオ評価

パフォーマンス評価

真正の評価

客観テスト

ワークシート

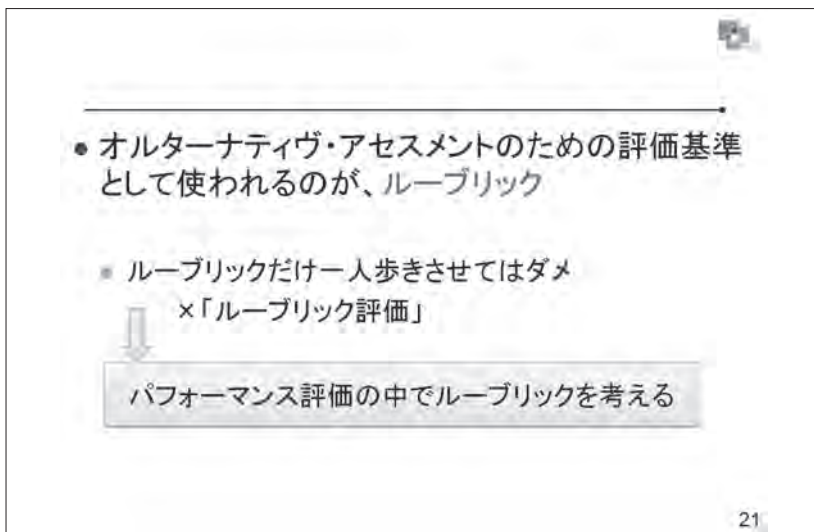
...

学習の証拠(エビデンス)となるようなもの

(cf. Wiggins & McTighe, 2005)

20

【スライド①- 21】



● オルターナティブ・アセスメントのための評価基準として使われるのが、ルーブリック

- ルーブリックだけ一人歩きさせてはダメ
×「ルーブリック評価」

↓

パフォーマンス評価の中でルーブリックを考える

21

【スライド①- 22】



パフォーマンス評価とルーブリック

22

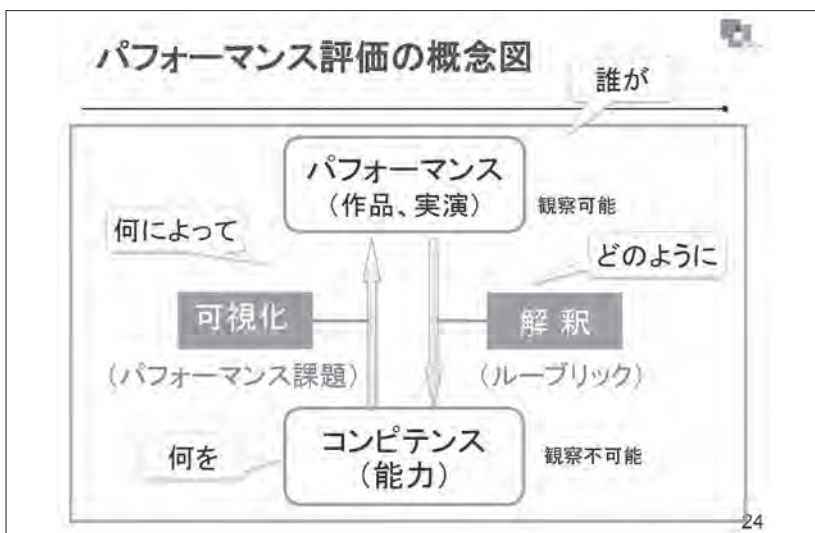
【スライド①- 23】

パフォーマンス評価についての問い

<ul style="list-style-type: none"> ● 誰が ● 何を ● 何によって (何をエビデンスとして) ● どのように 評価するか？ 	<ul style="list-style-type: none"> → ● 教員／学生が → ● 学習成果を → ● パフォーマンスによって → ● ルーブリックを使って 評価する
--	---

23

【スライド①- 24】



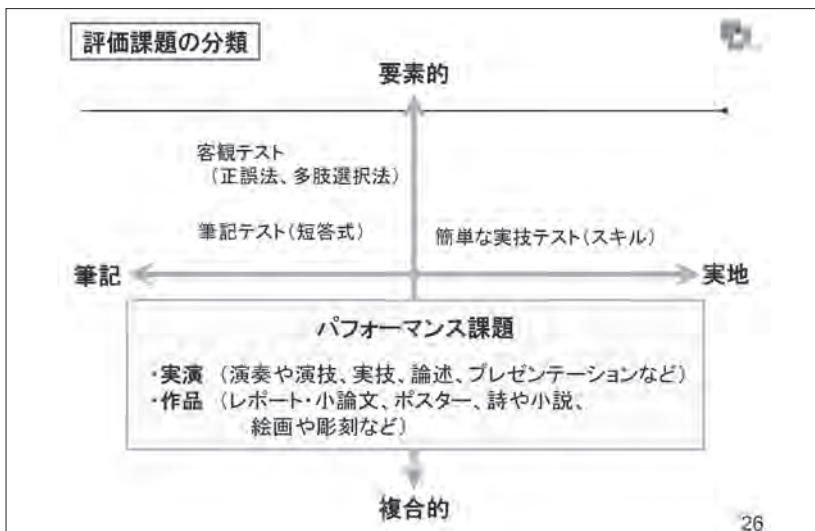
【スライド①- 25】

パフォーマンス課題

- パフォーマンス課題 (performance task)
 - 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課題
 - 様々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
「簡単にパフォーマンス評価の意図をいえば、テストの内容を、基準となるパフォーマンスで示される批判的な思考や知識の総合を求めるものにしていこうとするところにある」(ギブス, 2001, p.16)
- 大学教育の現状
 - パフォーマンス課題は多く使われている
実技(医療・教員養成など)、演奏や美術作品(芸術分野)、製作物(PBLなど)、レポート・卒業論文、プレゼン・口頭試問など
 - しかし、評価基準については、ほとんど主観にゆだねられていた

25

【スライド①- 26】



【スライド①- 27】

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
 - 中教審「質的転換答申」で言及
 - ← 厳格な成績評価、アカウントビリティの要請

	レベル4	レベル3	レベル2	レベル1	← レベル
観点1	
観点2	
観点3	
観点4	← 記述語 (descriptor)

27

【スライド①- 28】

ルーブリックのタイプ

- (a) 構造
 - 観点を複数設定して分析的に評価する「分析的(観点別)ルーブリック」
 - 観点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」
- (b) スコープ(適用範囲)
 - ある領域で一般的に適用できる「一般的ルーブリック」
 - 当該課題だけに適用される「課題特殊的ルーブリック」
- (c) スパン(対象期間)
 - 複数年にまたがって使われる「長期的ルーブリック」
 - 採点のためにスナップショット的に使われる「採点用ルーブリック」

28

【スライド①- 29】

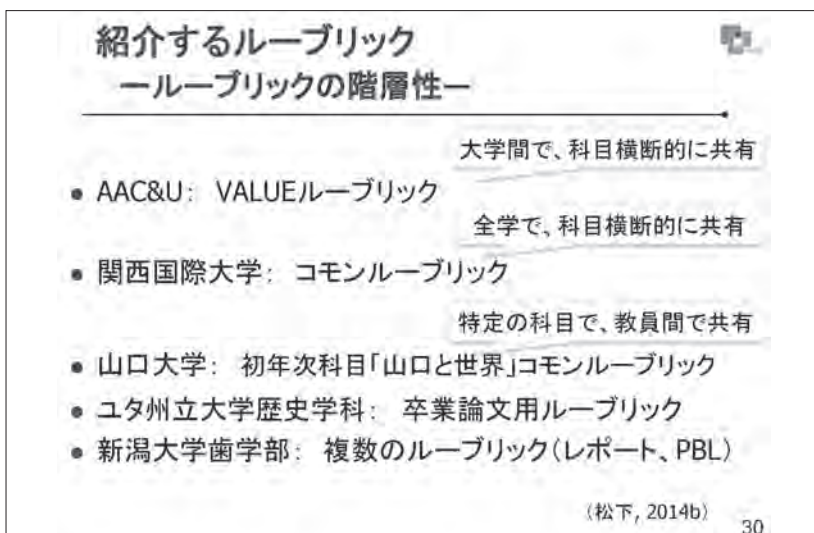


A slide with a white background and a thin black border. The title "ルーブリックの事例" is centered in a large, bold, black font. In the top right corner, there is a small, faint logo consisting of several overlapping squares. In the bottom right corner, the number "29" is displayed.

ルーブリックの事例

29

【スライド①- 30】



A slide with a white background and a thin black border. The title "紹介するルーブリック" is at the top left, and "ルーブリックの階層性" is below it, both in bold black font. A horizontal line with a right-pointing arrow is positioned below the subtitle. On the right side, there is a small, faint logo. The main content is a bulleted list of examples, each with a corresponding note to its right. At the bottom right, there is a citation "(松下, 2014b)" and the number "30".

紹介するルーブリック

ルーブリックの階層性

- AAC&U: VALUEルーブリック 大学間で、科目横断的に共有
- 関西国際大学: コモンルーブリック 全学で、科目横断的に共有
- 山口大学: 初年次科目「山口と世界」コモンルーブリック 特定の科目で、教員間で共有
- ユタ州立大学歴史学科: 卒業論文用ルーブリック
- 新潟大学歯学部: 複数のルーブリック(レポート、PBL)

(松下, 2014b) 30

【スライド①- 31】

例1. AAC&UのVALUEルーブリック

- VALUEルーブリック
 - もともとは15領域で共同開発（+1領域: グローバル学習）
 - 批判的思考／文章コミュニケーション／チームワーク／倫理的推論 など
 - 学士課程4年間の長期的ルーブリック（Rhodes, 2010; 松下, 2012）
 - 「メタルーブリック」としての性格（Rhodes & Finley, 2013）
 - …共通性と多様性の両立を図る

メタルーブリック

VALUE
ルーブリック

ローカライズ
(modification)

各大学・学科等の
ルーブリック

共同開発

さまざまな
ルーブリック

31

【スライド①- 32】

【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	マイルストーン		ベンチマーク
	4	3	2	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、作品全体を形づけている。	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まく実行している。	基本的構成や提示のしかたについて一貫した体系を使おうとしている。
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。	アイデアを裏づけるために、資料を使おうとしている。
撰文と技法を 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。ほとんど全く誤りがない。	用語法に誤りがあるために、意味の伝達が妨げられるような言葉遣いをしてい。

（注）レベル1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。

32

【スライド①- 33】

【例】ある大学の1年次科目(BIO109)の
レポート評価用ルーブリック

- 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 明確に伝えていた。	

(Rhodes, 2013) 33

【スライド①- 34】

例2. 関西国際大学のコモンルーブリック

- VALUEルーブリックを参考に、
 全学共通のコモンルーブリックを開発 (濱名他, 2012)
 - 3つのスキル: ライティング、プレゼンテーション、リサーチ
 - 全体版(6段階)ー下位学年版(4段階)ー上位学年版(4段階)
の長期的ルーブリック
 - 試行を重ね2~3年かけて作成、全授業の約6割で使用

34

【スライド①- 35】

● コモンルーブリック(ライティング)

共通ルーブリック「ライティング」2012年**版
共通ルーブリック「ライティング」2012年**版

	1	2	3	4	5	
議題に対する反応	議題に対する反応がほとんど見られない。議題について意見を述べていない。	議題に対する反応は若干見られるが、議題について意見を述べていない。議題について意見を述べているが、意見が不明確である。	議題に対する反応が若干見られるが、議題について意見を述べていない。議題について意見を述べているが、意見が不明確である。	議題に対する反応が若干見られるが、議題について意見を述べていない。議題について意見を述べているが、意見が不明確である。	議題に対する反応が若干見られるが、議題について意見を述べていない。議題について意見を述べているが、意見が不明確である。	議題に対する反応が若干見られるが、議題について意見を述べていない。議題について意見を述べているが、意見が不明確である。
論理の展開	論理の展開がほとんど見られない。論理の展開がほとんど見られない。	論理の展開がほとんど見られない。論理の展開がほとんど見られない。	論理の展開がほとんど見られない。論理の展開がほとんど見られない。	論理の展開がほとんど見られない。論理の展開がほとんど見られない。	論理の展開がほとんど見られない。論理の展開がほとんど見られない。	論理の展開がほとんど見られない。論理の展開がほとんど見られない。
レファレンス(引用)の適切性	レファレンス(引用)がほとんど見られない。レファレンス(引用)がほとんど見られない。	レファレンス(引用)がほとんど見られない。レファレンス(引用)がほとんど見られない。	レファレンス(引用)がほとんど見られない。レファレンス(引用)がほとんど見られない。	レファレンス(引用)がほとんど見られない。レファレンス(引用)がほとんど見られない。	レファレンス(引用)がほとんど見られない。レファレンス(引用)がほとんど見られない。	レファレンス(引用)がほとんど見られない。レファレンス(引用)がほとんど見られない。
文脈の考慮	文脈の考慮がほとんど見られない。文脈の考慮がほとんど見られない。	文脈の考慮がほとんど見られない。文脈の考慮がほとんど見られない。	文脈の考慮がほとんど見られない。文脈の考慮がほとんど見られない。	文脈の考慮がほとんど見られない。文脈の考慮がほとんど見られない。	文脈の考慮がほとんど見られない。文脈の考慮がほとんど見られない。	文脈の考慮がほとんど見られない。文脈の考慮がほとんど見られない。
表現の豊かさ	表現の豊かさがほとんど見られない。表現の豊かさがほとんど見られない。	表現の豊かさがほとんど見られない。表現の豊かさがほとんど見られない。	表現の豊かさがほとんど見られない。表現の豊かさがほとんど見られない。	表現の豊かさがほとんど見られない。表現の豊かさがほとんど見られない。	表現の豊かさがほとんど見られない。表現の豊かさがほとんど見られない。	表現の豊かさがほとんど見られない。表現の豊かさがほとんど見られない。
文法・句法	文法・句法がほとんど見られない。文法・句法がほとんど見られない。	文法・句法がほとんど見られない。文法・句法がほとんど見られない。	文法・句法がほとんど見られない。文法・句法がほとんど見られない。	文法・句法がほとんど見られない。文法・句法がほとんど見られない。	文法・句法がほとんど見られない。文法・句法がほとんど見られない。	文法・句法がほとんど見られない。文法・句法がほとんど見られない。
読者の意識	読者の意識がほとんど見られない。読者の意識がほとんど見られない。	読者の意識がほとんど見られない。読者の意識がほとんど見られない。	読者の意識がほとんど見られない。読者の意識がほとんど見られない。	読者の意識がほとんど見られない。読者の意識がほとんど見られない。	読者の意識がほとんど見られない。読者の意識がほとんど見られない。	読者の意識がほとんど見られない。読者の意識がほとんど見られない。

上位学年版 下位学年版

35

【スライド①- 36】

例3-1. 山口大学の共通ルーブリック

● 全学必修の初年次科目「山口と世界」のための共通ルーブリックを開発

- プロトタイプとしてのルーブリック
(同じ「共通ルーブリック」でも意味が異なる)
- 個々の教員が自分の担当する「山口と世界」にあわせて修正できる(modification) * 調査、研究、商品開発など
- 成績評価にも使えるようになっている

36

例3-2. ユタ州立大学歴史学科の 卒業論文用ルーブリック

* ユタ州立大学歴史学科の取組は、歴史学の分野で、ユタ州内の全
大学、全米(アメリカ歴史学会)にも広がっている

- 歴史学科全体で、学生が身につけるべき学習成果を設定
 - 歴史学の知識／歴史学の思考／歴史学のスキル
- 最終学年でそれが獲得されているかを、卒業論文で評価するためのルーブリック
 - 成績評価にも使えるようになっている

山口 隆雄

37

例3-3. 新潟大学歯学部での パフォーマンス評価の開発

- 初年次科目「大学学習法」のレポート評価用の
ルーブリック開発 (松下・小野・高橋, 2013)
 - パフォーマンス課題はレポート
 - ルーブリックの開発
- 2～5年次の専門科目「PBLテュートリアル」のための
パフォーマンス評価開発 (小野・松下・斎藤, 2014)
 - パフォーマンス課題の開発
 - ルーブリックの開発



38

第 2 部

「大学学習法」の学習評価実践

—レポート評価に焦点をあてて—



新潟大学 歯学部 教授

小野 和宏 氏

はじめに

新潟大学歯学部の小野と申します。このような機会を与えていただき、心からお礼を申し上げます。私は松下先生のお話を受け、パフォーマンス評価の実践事例ということで、「大学学習法」の学習評価実践についてお話しさせていただきます。松下先生が、最後に「苦労話」とおっしゃっていましたが、なかなかうまくいかないことも多く、そこも含めてお話しさせていただこうと思っています。

最初に、新潟大学歯学部についてご紹介します。新潟大学歯学部は、歯科医師を養成する6年制の歯学科と、口腔生命福祉学科という歯科衛生士と社会福祉士の能力をあわせもつ人材を養成する4年制の学科の2学科からなっています。それぞれここに書いてあるように、1学年の学生定員が40名、20名と、比較的小さい医療系の学部になります【スライド②-2】。立教大学とはさまざまな点で状況が異なるでしょうが、私たちの取組みが何かしらお役に立てば幸いです。

1. 学士課程をどうみるか

私たちは2001年から、学士課程教育に関して、ここに書いてある基本的認識を共有しています【スライド②-3】。「学士課程教育を歯科医療従事者としての生涯学習の最初の段階と位置づけ、問題解決能力の育成を重視し、その後にくる大学院や実社会での学習のなかで、専門性を主体的に向上させる人材を養成する」ということです。このなかでキーワードは、「問題解決能力」ということ

になるのですが、こういう考えにいたったのは、この漫画に描かれているような状況があったからです【スライド②-3】。これは医学教育を風刺した漫画です。教授が学生に、次から次へと知識を詰め込んでいる。学生のほうは、そんなに食べられないし、ましてや消化、吸収して自分のものにできるかというところと全然できないという状況です。

私たち教員としては、この教授の気持ちも分かりますよね。ここにいらっしゃる先生のどの専門分野でもそうだと思うのですが、今の世の中、日々新しい知見が生まれてくるわけです。そうすると、学生が社会に出て仕事についてのことを考えて、どうしても学生にたくさん教えたいという気持ちになります。ですが、この漫画が示しているように、実際には、学生がそれらすべてを消化するのはとても無理だと。そうであれば、自分には何が必要かわかって、自分で学ぶことができる、そういう能力をつけてあげたほうが、これからの学生は伸びるだろう。このような認識から、学士課程教育をとらえ直し、「学士課程では学生に自立・自律した学習者としての基礎を着実に身につけさせる」という考えにいたっているわけです。今日、過分なお紹介をしていただきましたが、歯学部でレポート評価の取組みって何か奇異に感じませんか。歯学部なら「レポートの評価より、たとえば、歯をうまく削る練習をしたほうがいいのでは」と考えますよね。今のお話で分かっていただけたと思いますが、学士課程に対する私たちの課題認識が



背景にあって、これからご紹介するレポート評価に取り組んできたということです。

2. 「大学学習法」という科目

本題の大学学習法の話に入ります。この科目では初年次前期にこのような内容を学びます【スライド②-4】。どこの大学でもやられている内容だと思います。どういことを科目の目標にしているかという、今ほどお話ししたように、「自ら問題を見つけ、必要な情報を収集、分析、統合し、論理的に思考して問題を解決する基本的な能力を身につける」【スライド②-5】。すなわち、学生は、講義や演習で身につけた知識や技能を使って、プレゼンテーションをして、クラスメートや教員とのディスカッションを参考にして、最終的にレポートを提出する。そして、この学習活動のなかで、今言ったような能力を育成していこうという授業です。

3. パフォーマンス評価への着目

ところが、松下先生のお話にもありましたように、そういう能力を、どう評価するのかということが非常に難しい。どう評価してよいか分からず、これまでは定型的な文書作成や、口頭発表のスキルみたいな、要するにテクニカルなことを評価してきました。すなわち、能力目標と評価の間に齟齬がありました。こういうことをやっていると、「シラバスにはこう書いてあるけれども、先生はもう少し違ったところをみているのかな、大事にしているのかな」ということで、学生にあまりよろしくないメッセージを暗黙のうちに出示してしまいます。また、学習に対するモチベーションをそぐような形にもなっていたかもしれないという反省がありました。それだったら、ちょっと大変かもしれないけれど、レポートから問題解決能力、論理的思考力、表現力を把握する、妥当性、信頼性、実行可能性の高い評価方法を何とか開発してみようということになりました。そのようにしてこの取組みが始まり、そこではじめてパフォーマンス評価に注目したということになります。

今までお話ししてきた内容をまとめたものがこれです【スライド②-6】。上は

大学学習法の授業の構造で、下は先ほど松下先生がご紹介くださったパフォーマンス評価の概念図です。それを上下に並べました。大学学習法は、ここにあるようなことを講義・演習で教え、その知識・技能を使って、プレゼンテーションをさせ、またレポートを書く学習活動をさせます。そのなかで、先ほど述べた教育目標に到達させようという形になっています。一方、パフォーマンス評価の手法というのは基本的に同じような形で進みます。知識・技能を学ばせて、何かしらパフォーマンスをさせ、そのなかから、自分たちが目標としていた能力がどのくらい身についたかを把握する。このような構造をもつパフォーマンス評価であれば、大学学習法の教育目標とした能力を把握できるだろうと考えたわけです。

大学学習法の課題は、プレゼンテーションとレポートですが、先ほどお話ししたように、まずプレゼンテーションして、そこでのディスカッションを踏まえて、最終的にレポートを出すものですから、パフォーマンス課題はレポートということにしました。

4. ルーブリックの開発を始める

そうすると残るのがルーブリックで、その開発を始めたということになります。これが2011年度に、私たち歯学部の教員7名が集まって考えたものです【スライド②-7】。便宜上、今回の発表ではバージョン1とさせていただきます。「知識・理解」「問題発見」「情報検索」「論理的思考と問題解決」「文章表現」という5つの観点があり、レポートの構成要素を網羅しています。

しかし、「問題解決能力」や「論理的思考力」と一言で言うのですが、把握したいこれらの能力がどのような能力であるかという定義の検討が不十分で、能力と観点の関連が不明確です。どの観点のレベルが高ければ、どの能力が高いのか、そういうことがよく分からない。あと、ここに「知識・理解」が入っています。他所でみるルーブリックの例でも入っていることが多いのですが、そもそもルーブリックを作り始めたときに何を評価したかったのかということが、みんなで話し合っているうちに頭から抜けてしまったのですね。レポートの内容、すなわち知識の量や意見の当否を評価することが本来の目的ではなかったはずなのに、ここに入り込んできたと。

それから、「論理的思考と問題解決」は、カテゴリーが大き過ぎて、教育的利

用という点で問題でした。どういうことかということ、カテゴリーが大きいので、レベルが高かったり低かったりしたときに、学生のどこがよくて、どこが悪いのかということ、指導に結びつけられないということです。先ほどの松下先生のお話のなかに、「全体的ルーブリック」「ホリスティックなルーブリック」という話がありましたが、それと同じことがここで起こってきました。

15. ルーブリックの改訂 1 ―観点を見直す―

「三人寄れば文殊の知恵」と言いますが、どうも歯学部教員7人集まってもよい考えは浮かばないように、このバージョン1の作成を契機に、松下先生との共同取組みが始まりました。「教育の専門家と一緒にやったほうがいいよね」ということです。

最初に問題解決能力や論理的思考力を定義し直そう、というか、定義しようということになりました。そこで、関西大学の牧野先生の「論理のしくみ図」（牧野 2008, p.207）と「論理のしくみ図シート」（牧野 2008, p.219）を参考にさせていただき、そこから問題解決と論理的思考力というものを定義し、もう一度ルーブリックを作り直すことにしました。

これが2012年度の授業で使ったルーブリックで、バージョン2ということ



になります【スライド②-8】。「背景と問題」「意見と結論」「根拠・情報」「異なる意見の検討」「全体構成」「表現ルール」の6つの観点にしました。観点の上段では、「背景と問題」「意見と結論」は「この2つの観点では問題解決をみているのですよ」ということを示しています。同様に、「意見と結論」「根拠・情報」「異なる意見の検討」「全体構成」では、論理的思考をみようとしている、ということを表しています。「全体構成」と「表現ルール」の観点は文章表現をみるものです。このような形で、6観点のルーブリックを作りました。

ちなみに後の話に関係してきますので、レベルの話をしておきます。レベルはバージョン1もバージョン2もレベル1からレベル3までの3段階になっています。レベル1を読んでもらうと分かると思いますが、「ほとんど何もできていない」というようなことが書いてあります【スライド②-7、②-8】。

16. 評価の信頼性の検討

このルーブリックを用いて、2012年度はレポート64本を、教員3名で評価しました。このときのレポート課題は松下先生に「あまり好ましくない」と言われているのですが、一貫して社会問題をレポート課題にしていたものですから、当時、「エネルギー問題」と「大学の秋入学問題」が話題でしたので、それにし



ました。なお、「社会問題は好ましくない」というのは、後の話にも関係しますが、いろいろな意見がインターネット上に溢れていて、学生がコピペに陥りやすく、大学学習法が目標としている能力を育成できないのではないかとこの危惧からです。

評価してみて、その評価の信頼性がどの程度なのか検討しようということになりました。一緒に共同研究をしている統計に詳しい先生から、「最初だから、一般化可能性理論を使って、どんなルーブリックにすれば信頼性が上がるのかといった点も含めて検討するのがいいのでは」との助言をいただき、そのような検討を行いました。

まず一般化可能性係数は 0.62 で、通常、0.8 で十分な信頼性といわれているのですが、そこまでは達していないことが分かりました【スライド②-9左の表】。ただ、先行研究と比べると、必ずしも低い値ではなく、ほぼ同じ値ですので、そこそこの信頼性はあったということになります。信頼性が低くなった理由は、ここに書いてありますが、学生の分散成分の全体に占める割合が 13.87% と非常に低い。このことが一番の原因だと考えられます。要するに学生個々の得点のばらつきが、レポートの質のばらつきとパラレルになっていないという、うまく説明できていないかもしれませんが、学生のレポートの質を得点が的確に表していないことが原因だということです。先ほどちょっとお話しましたが、レベルが 3 つしかなく、かつレベル 1 が非常に低い形、逆にレベル 3 が非常に高い形に設定されていますので、レベル 2 の幅がかなり広がっています。そういうレベル設定になっていたことが関係しているだろうと考えました。

それから、評価の観点や評価者の数を増やすと信頼性はどうなるのかというシミュレーションを行いました【スライド②-9右の図】。観点は、すでに 6 つですから、6 つから増やしても、あまり信頼性は変わらない。あと、評価者を増やせば、このときは 3 人ですので、ぐっと信頼性は上がるのですが、「実行可能性という点からは現実的ではないよね」という結論で、少しまたルーブリックを変えてみようということになりました。

7. ルーブリックの改訂 2

—信頼性分析を踏まえレベル数を増やす—

それが、バージョン3という名前にさせていただいたもので、2013年度、2014年度と使っています【スライド②-10】。観点については、多少文言が変わっていますが基本的にバージョン2と同じです。レベルについては4段階にしました。すなわち、レベル1を上げて、レベル1に満たないものをレベル0とすることで4段階にしました。

これを使って、2013年度はレポート60本を教員3名で、2014年度はレポート63本を教員4名で評価しました。レポート課題はこの2年間は「医療や科学の進歩をもたらした諸問題」ということにしています。

お配りしている別紙は、今年度、実際に学生に配った課題シートです（【別紙資料】p.80参照）。一番上に課題が書いてあって、その下に注意事項が書いてあります。ルーブリックも学生に渡しました。課題シートの一番下のところに書いてありますが、「ルーブリックを見て、提出する前にレポートを自己評価してね、その自己評価で自分はどの能力がよかったとか、悪かったとかを、自分で考えて、今後の学習の役に立ててね」というようなことが書いてあります。

8. レベル数の設定と評価者間信頼性

2012年度から2014年度までのレポート評価に関する分析結果をここに一覧としてまとめました【スライド②-11】。学生数と、評価者数は上の表でお示ししているとおりです。評価者は同じで、今年度だけ、新しい評価者が1名加わっています。レポート課題、使用したルーブリックのバージョンと特徴もご覧のとおりです。下の表は、レポート評価の結果について、級内相関係数で評価者間信頼性を調べ、それを各年度の、各観点についてまとめています。

年度間で統計学的に有意差があるわけではないので、あくまでも誤差の範囲内で数値の上がり下がり問題をしていると言われればそのとおりなのですが、それを分かったうえで、この数値が上がったり、下がったりしている意味を考察してみたいと考えまして、これから少しお話しさせていただきます。

まず2012年度と2013年度では、ルーブリックが変わりました。大きな変

更点として、3レベルから4レベルに変わっています。そうすると、当初予想したように、若干ではあるのですが、全ての観点で評価者間信頼性が上がっています。もちろんレポート課題も違いますし、統計学的有意差があるわけではないので、はっきりしたことは言えないのですが、レベル数が3よりは4のほうが、信頼性という点ではよかったのではないかとということがうかがえます。ルーブリックのレベルを何段階に設定するかという問いが、どれくらい意味のある問いかわかりませんが、あまり少な過ぎると、先ほどのバージョン2で述べたようなことが起こります。また、私は6レベルや8レベルのルーブリックで評価した経験がないので何とも言えないのですが、あまりに多すぎると、レベル間のパフォーマンスの違いがはっきりせず、どのレベルにするか評価者はかなり判断に迷うのではないかという気がします。ですので、4段階くらいが適当ではないかなと、いかげんな意見として聞き流していただければよいのですが、そういうこともこの2012年度と2013年度の差異から言えるかもしれません。

19. 評価者によって異なる判断基準

次に、2013年度と2014年度をみてみましょう。2014年度は4名で評価しましたが、これまでやってきた3名でも値を出していますので【スライド②-11】、「2013年度」と「2014年 従来評価者3名」とは、学生が違うことを除けば、レポート課題、ルーブリック、評価者は全く一緒の条件です。比べてみると、評価者間信頼性は、観点により上がったり、下がったりしています。

評価を繰り返していけば、信頼性は上がってくると思うのですが、そう簡単にはいかないようです。上がっていない観点はどれかということ、「主張と結論」と「根拠と事実・データ」、それから「表現ルール」です。これらの観点の特性を踏まえると、コピペが疑われる場合の取り扱いや、学生が提示した事実・データが信頼できるものであるかという判断に関する不一致に起因しているのではないかと考えられます。ルーブリックを見てください【スライド②-10】。「主張と結論」のところには、「自分の」と書いてあります。すなわち、コピペが疑われる主張や結論をどう扱うか、そもそもコピペを教員は見破れるか、評価者によって違いがあるのでしょうか。「表現ルール」のなかには、「引用部分と自分の文章の区別を明示し」と書いてあるので、「主張と結論」と同じように、コピペして引用元を

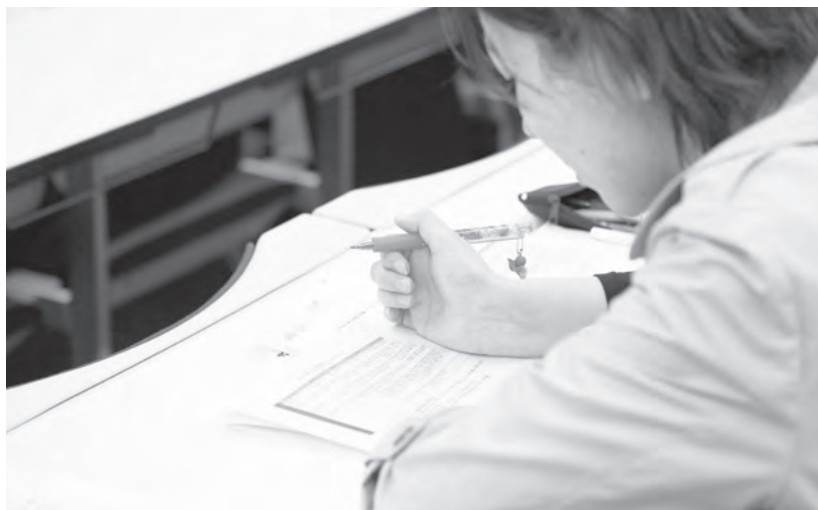
示していない部分を見破れるか評価者にかかってきます。「根拠と事実・データ」では、学生は当然いろいろなリソースを使いますよね。たとえば、ネットから得た情報。それを信じていいのか悪いのか、評価者は判断しなければなりません。

実際には、3名で評価しているうちの1名が特異な評価をしているのですね。トレーニングも必要でしょうし、よく言われることですが、評価の基準合わせは極めて重要だということを、おそらくこのデータは示しているのだと思います。

なお、先ほど一般化可能性理論を用いてシミュレーションしたように、「2013年度」と「2014年度 評価者4名」のところを見ていただくと分かりますが【スライド②-11】、評価者が3名から4名になると信頼性は上がっています。

10. 締めくくり

ほとんど推論に推論を重ねたような話で、申し訳ありませんでした。これで私の話は大体終わりなのですが、いままで丸3年ほど取り組んできました。一つよかったことを挙げれば、ループリックによって、問題解決能力、論理的思考力、表現力といった、教員もよく分からないし、教えられる学生もよく分からないという状態だった曖昧な概念を、学習や指導の場面で理解可能なレベルまで、具体化することに成功したということではないかと思えます【スライド②-12】。



それと、レポートから高次の統合的な能力を、妥当性と一定の信頼性をもって把握することができたといえるのではないのでしょうか。信頼性を向上させることにより、教授・学習にとってより効果的なツールとなるよう、今後さらに検討が必要ですが。

信頼性にもいろいろあります。なぜ、私たちが評価者間信頼性を重視しているかということ、この評価者間信頼性を高めることによって、実行可能性も高まるのではないかと考えているからです。すなわち、何人かの教員が評価して、同じような結果が出るのであれば、レポートを分担して評価することもできるかなと。いま、1人の教員が学生レポート60本を評価していますが、正直なところ、評価負担は感じています。「レポート60本で弱音を吐くな」という声が聞こえてきそうですが、しかし評価者間信頼性が上がってくれば、分担しても信頼性ある評価ができますから、1人で60本を評価する必要がなくなるかもしれません。このように実行可能性とも関連してくると考えて、この指標をみているということですね。

課題は、実はまだまだたくさんあって、今日は学習評価実践としてお話しさせていただきましたが、本当の意味での学習評価実践にするには、ここで終わってはだめだと思うのです。これまでは、言ってみれば、ただ単に評価方法を開発してきただけで、評価した結果を学生にフィードバックして、学生の能力をさらに



伸ばしていくとか、これから松下先生のお話にあると思いますが、学生が自己評価の主体となるように学生を伸ばしていく、そういうところにまでもっていくことができ、はじめて学習評価実践というものになるのだと思います。

うまくご紹介、ご説明できたか自信がありませんが、これで私の事例報告を終わりにさせていただきます。早口で、またお聞き苦しい点もあったと思いますが、どうぞお許してください。ご清聴ありがとうございました。

【スライド②-1】

立教大学教育開発・支援センター主催 秋期シンポジウム (2014年10月23日 立教大学)
「学習成果」の設定と評価—アカデミック・スキルの育成を手がかりに—

「大学学習法」の教育評価実践
—レポート評価に焦点をあてて—



新潟大学歯学部
小野和宏

k-ono@dent.niigata-u.ac.jp

【スライド②-2】

新潟大学歯学部

歯学科(1965年～)

教育課程:6年制

歯学教育プログラム

学 位:学士(歯学)

資 格:歯科医師国家試験受験資格

学 生 数:40名/学年(+3年次編入学生5名)

口腔生命福祉学科(2004年～)

教育課程:4年制

口腔保健福祉学教育プログラム

学 位:学士(口腔保健福祉学)

資 格:歯科衛生士国家試験受験資格

社会福祉士国家試験受験資格

学 生 数:20名/学年(+3年次編入学生6名)



【スライド②-3】

学士課程教育に対する基本的認識 (2001～)

学士課程教育を歯科医療従事者としての生涯学習の最初の段階と位置づけ、問題解決能力の育成を重視し、その後に続く大学院や実社会での学習のなかで、専門性を主体的に向上させうる人材を養成する



学士課程では学生に自立・自律した学習者としての基礎を着実に身につけさせる



【スライド②-4】



大学学習法

開講：初年次前期15週30回

対象：歯学科・口腔生命福祉学科1年生(60名)

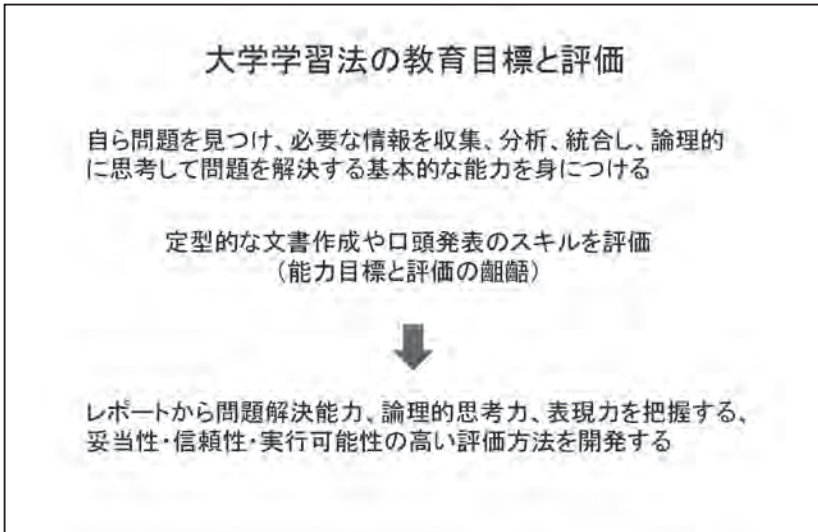
授業(自作テキストを使用した講義・演習)：

- ・ 教育プログラムでの大学学習法の位置づけ
- ・ 図書館の利用方法
- ・ コンピュータの基本操作と電子メールの利用
- ・ インターネットの利用とセキュリティ
- ・ 情報検索の方法
- ・ ウェブページの作成と画像の取り扱い
- ・ ワード・エクセル・パワーポイントの基本操作
- ・ レポートの書き方
- ・ プレゼンテーションの仕方

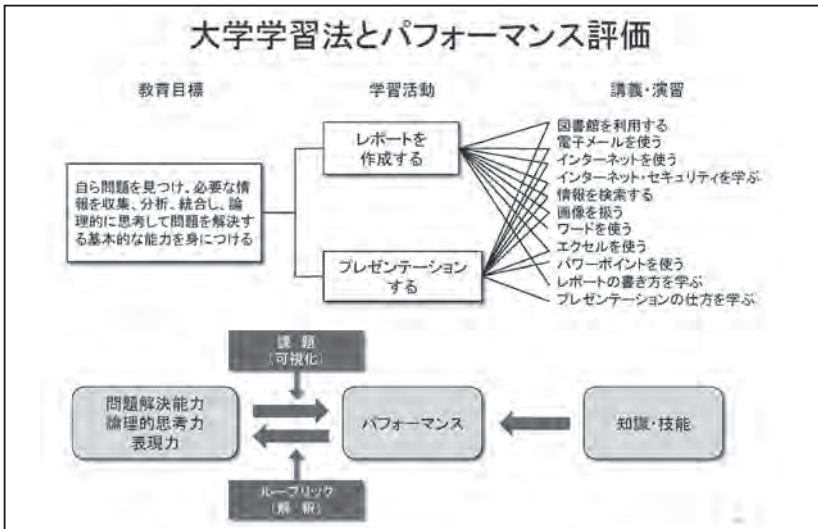
課題(評価対象)：

- ・ プレゼンテーション(10分)・質疑応答
- ・ レポート(概要600字 本文3200字)

【スライド②-5】



【スライド②-6】



[スライド②-7]

ルーブリック2011 (Ver. 1)

観点	知識・理解	問題発見	問題解決	論理的思考と問題解決	文章表現
総合的	自身の取組する課題領域、さまざまな観点から検討していることを示す。	与えられたテーマに適切に取り組むための、適切な問題を発見する。	問題解決の必要とする情報を収集し、その活用を検討する。	収集した資料の中から情報をもとに、適切な見方を導き出す。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現する。
1-1	自分の取組する課題について、多面的・多角的に検討し、その本質を掘り下げる。	与えられたテーマから自分の課題を発見する。その問題を取り上げた理由や背景について述べている。	自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用している。	収集した情報や資料を適切にまとめ、自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用している。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現する。
1-2	自分の取組する課題について、多面的・多角的に検討しているが、その本質の掘り下げが不十分である。	与えられたテーマから自分の課題を発見しているが、その理由や背景について述べていない。	自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	収集した情報や資料を適切にまとめ、自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現するが、その本質の掘り下げが不十分である。
1-3	自分の取組する課題について、一面での見解にとどまっている。	与えられたテーマから自分の課題を発見しているが、その理由や背景について述べていない。	自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	収集した情報や資料を適切にまとめ、自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現するが、その本質の掘り下げが不十分である。
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

- 5つの観点を持ち、レポートの構成要素を網羅
- 把握したい問題解決能力や論理的思考力とどのような能力であるかという定義の検討が不十分で、能力と観点の関連が不明確
- レポートの内容(知識の量や意見の当否)を評価することが目的ではない
- 「論理的思考と問題解決」はカテゴリが大きすぎ、教育的利用という点で問題

[スライド②-8]

ルーブリック2012 (Ver. 2)

観点	知識・理解	問題発見	問題解決	論理的思考	文章表現
総合的	自身の取組する課題領域、さまざまな観点から検討していることを示す。	与えられたテーマに適切に取り組むための、適切な問題を発見する。	問題解決の必要とする情報を収集し、その活用を検討する。	収集した資料の中から情報をもとに、適切な見方を導き出す。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現する。
1-1	自分の取組する課題について、多面的・多角的に検討し、その本質を掘り下げる。	与えられたテーマから自分の課題を発見する。その問題を取り上げた理由や背景について述べている。	自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用している。	収集した情報や資料を適切にまとめ、自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用している。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現する。
1-2	自分の取組する課題について、多面的・多角的に検討しているが、その本質の掘り下げが不十分である。	与えられたテーマから自分の課題を発見しているが、その理由や背景について述べていない。	自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	収集した情報や資料を適切にまとめ、自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現するが、その本質の掘り下げが不十分である。
1-3	自分の取組する課題について、一面での見解にとどまっている。	与えられたテーマから自分の課題を発見しているが、その理由や背景について述べていない。	自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	収集した情報や資料を適切にまとめ、自分の課題を捉えようとする中で、その問題の背景や、レポートの作成に必要とする情報を収集し、活用しているが、その本質の掘り下げが不十分である。	採集した資料を適切に整理し、論理的にまとめる過程を、適切に表現するが、その本質の掘り下げが不十分である。
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

- 学生レポート64本を教員3名が評価
- レポート課題は「エネルギー問題」「大学の秋入学問題」

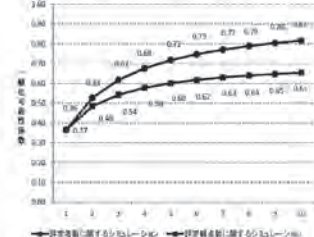
[スライド②-9]

一般化可能性理論にもとづく方法による レポート評価の信頼性の検討

各分散成分の推定値と一般化可能性係数

要因	分散成分 の推定値	(%)
学生 (p)	0.065	13.87%
評価観点 (r)	0.054	11.70%
評定者 (v)	0.042	9.09%
学生 × 評価観点 (p × r)	0.022	4.65%
学生 × 評定者 (p × v)	0.077	16.47%
評定者 × 評価観点 (r × v)	0.010	2.07%
学生 × 評定者 × 評価観点 (p × r × v)	0.196	42.15%
一般化可能性係数	0.62	

一般化可能性係数のシミュレーション結果



- ・ 学生の得点の分散成分 (p) が全体に占める割合は13.87%と低い
- ・ 評価観点を増やしても信頼性はあまり高まらない
- ・ 評価者数を増やすのは現実的ではない

[スライド②-10]

ループリック2013 (Ver. 3)

観点	内容と特徴	立場と認識	理解の進捗	評価の進捗	非身体的に	文章表現
目的意識	各人自らのテーマから自分で問題を発見する。	既知した問題に対し、既知してきた事柄の整理と統合から、疑問を抱く。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	問題の整理と統合する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	結論やポイントとして、ゴール・構成要素を、問題と整合し、論理的に、論理のゴールを構成する。
問題発見	各人自らのテーマから問題を発見し、深まる過程も含め、その問題を解かした後の再考も含めて行うべきである。	既知した問題に対し、既知してきた事柄の整理と統合から、疑問を抱く。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	問題の整理と統合する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	結論やポイントとして、ゴール・構成要素を、問題と整合し、論理的に、論理のゴールを構成する。
問題解決	各人自らのテーマから問題を発見し、その問題意識を深めながら、論理を深めながら、問題を解かした後の再考も含めて行うべきである。	既知した問題に対し、既知してきた事柄の整理と統合から、疑問を抱く。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	問題の整理と統合する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	結論やポイントとして、ゴール・構成要素を、問題と整合し、論理的に、論理のゴールを構成する。
問題意識	各人自らのテーマから問題を発見し、その問題意識を深めながら、論理を深めながら、問題を解かした後の再考も含めて行うべきである。	既知した問題に対し、既知してきた事柄の整理と統合から、疑問を抱く。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	問題の整理と統合する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	結論やポイントとして、ゴール・構成要素を、問題と整合し、論理的に、論理のゴールを構成する。
問題解決	各人自らのテーマから問題を発見し、その問題意識を深めながら、論理を深めながら、問題を解かした後の再考も含めて行うべきである。	既知した問題に対し、既知してきた事柄の整理と統合から、疑問を抱く。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	問題の整理と統合する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	結論やポイントとして、ゴール・構成要素を、問題と整合し、論理的に、論理のゴールを構成する。
問題意識	各人自らのテーマから問題を発見し、その問題意識を深めながら、論理を深めながら、問題を解かした後の再考も含めて行うべきである。	既知した問題に対し、既知してきた事柄の整理と統合から、疑問を抱く。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	問題の整理と統合する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	結論やポイントとして、ゴール・構成要素を、問題と整合し、論理的に、論理のゴールを構成する。
問題解決	各人自らのテーマから問題を発見し、その問題意識を深めながら、論理を深めながら、問題を解かした後の再考も含めて行うべきである。	既知した問題に対し、既知してきた事柄の整理と統合から、疑問を抱く。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	自分の言葉で整理する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	問題の整理と統合する過程を通じて、問題の整理と統合する過程がテーマとして現れる。	結論やポイントとして、ゴール・構成要素を、問題と整合し、論理的に、論理のゴールを構成する。

- ・ 学生レポート60本を教員3名(2013)、63本を教員4名(2014)が評価
- ・ レポート課題は「医療や科学の進歩をもたらした諸問題」

【スライド②-11】

2012～2014年度におけるレポート評価

	学生	評価者	レポート課題	ルーブリック	ルーブリックの特徴
2012年度	64名	3名 (KO, HN, YA)	エボルキーン問題 大学の秋入学問題	Ver. 2	6観点3レベル 一時的・分析的・検査用
2013年度	60名	3名 (KO, HN, YA)	医療や科学の進歩がもたら した諸問題	Ver. 3	6観点4レベル 一時的・分析的・長期的
2014年度	63名	4名 (KO, HN, YA, JN)	医療や科学の進歩がもたら した諸問題	Ver. 3	6観点4レベル 一時的・分析的・長期的

級内相関係数による評価者間信頼性

	背景と問題	主張と結論	根拠と事実・ データ	対立意見の 検討	全体構成	表現ルール	計
2012年度	0.15	0.22	0.18	0.11	0.27	0.40	0.30
2013年度	0.15	0.30	0.33	0.44	0.30	0.45	0.51
2014年度 従来の報告3名	0.22	0.12	0.15	0.62	0.26	0.30	0.36
2014年度 新報告4名	0.21	0.36	0.35	0.67	0.36	0.53	0.56

- ・ ルーブリックのレベルの数を増やすし、各レベルの記述語をより学生の水準にそったものとすることで評価者間信頼性は高まった？
- ・ 評価者トレーニングにより評価の信頼性は高まる？
- ・ コピー&ペーストの取り扱い、信頼できる事実・データの判断に関する不一致？

13

【スライド②-12】

まとめと今後の課題

- ・ ルーブリックにより、「問題解決能力」「論理的思考力」「表現力」といった曖昧な概念を、学生の学習や教員の指導にとって理解可能なレベルにまで具体化することに成功した
- ・ レポートから、高次の統合的な能力を、妥当性と一定の信頼性をもって把握することができた
- ・ 信頼性を向上させることにより、教授・学習にとってより効果的なツールとなるよう検討が必要である

14

【スライド②-13】

参考文献

- ギブス,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。
- 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—』(田中精治監訳)ミネルヴァ書房。
- 牧野由香里(2008)『「議論」のデザイン—メッセージとメディアをつなぐカリキュラム』ひつじ書房。
- 松下佳代(2007)『パフォーマンス評価—子どもの思考と表現を評価する—』日本機軸。
- 松下佳代(2012)「パフォーマンス評価における学習の質の評価—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』18, pp.75-114。
- 松下佳代・小野和宏・高橋進介(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』35-1, pp.107-115。
- 小野和宏・八木稔・ステガロユ・ロクサーナ・大島勇人・西山秀昌・八巻正樹・前田健康(2011)「初年次教育としての大学学習法の効果と課題」『日本歯科医学教育学会雑誌』27, pp.69-77。
- 小野和宏・西山秀昌・八木稔・ステガロユ・ロクサーナ・重谷佳見・山村健介・井上誠・前田健康(2014)「大学学習法へのパフォーマンス評価導入における実践的課題」『新潟大学高等教育研究』1-2, pp.5-8。

【別紙資料】

2014年度 歯学スタディ・スキルズ 課題

医療や科学の進歩にはめざましいものがあります。少し前までは治療することが難しかった病気にも対処できるようになりました。また、便利な道具も世の中に溢れており、日々、私たちは恩恵を受けて暮らしています。しかしその一方で、進歩をもたらしたさまざまな問題も抱え込んでしまいました。皆さんもテレビや新聞で見聞きしていることでしょう。その解決のために、多くの議論がなされていますが、どの問題もしかるべき解決策が示されていないのが現状ではないでしょうか。

そこで、与えられたテーマ「医療や科学の進歩をもたらした諸問題」から話題を設定し、主体的に調査・学習を行い、自分の考えをレポートにして提出してください。

- * この用紙をレポートの表紙に添付して、7月25日午後4時30分までに歯学部学務係に提出してください。
- * レポートのワード・ファイルも小野 (kono@dent.niigata-u.ac.jp) に電子メールで提出してください。
- * レポートの書式は、テキスト 80 頁 (ワードの実習・ファイル提出用には下記の設定を使うこと) に準じてください。
- * レポートは、概要 400～600 字、本文 2800～3200 字としてください。
- * なぜこの話題を選んだのか、また話題に関する時代的・社会的背景など、わかりやすく記述してください。
- * どのような調査結果を得たか、またそれらの事実やさまざまな意見からどのような結論にいたったかを、配布したルーブリックを意識しながら、「なるほど!」と読者を納得させるよう論理的に記述してください。
- * 必要に応じて図表などを挿入してもかまいません (図表は字数に含めません)。
- * 書籍やウェブサイトから引用した箇所には、テキスト 112～113 頁にあるように肩番号をつけ、レポートの最後に引用文献を記載してください (引用文献は字数に含めません)。

例

「～」といわれている¹⁾。 「～」との意見もある²⁾。 「～」と報告されている³⁾。

引用文献

1) 新潟太郎: 少子高齢化と医療制度. 社会歯科学会雑誌, 23: 45-57, 2010.

2) 越後花子: 医療保険. 245-262 頁, 中央出版社, 東京, 2009.

3) 歯科医師臨床研修マッチング協議会: 歯科医師臨床研修マッチングプログラム説明資料. <http://www.drmp.jp/dmp.050813.pdf>

(閲覧日 2012 年 6 月 10 日)

自己評価

背景と問題	主張と結論	根拠と事実・データ	対立意見の検討	全体構成	表現ルール
■	■	■	■	■	■

- * 配布したルーブリックに従って、自分のレポートを評価し、■にレベルを記入してください。
- * 自己評価結果は成績とは関係ありません。過大評価・過小評価することなく、第三者の立場で自分自身のレポートを読んで、感じたまま素直に判断してください。
- * 自己評価結果から、今回の学習により「問題解決能力」「論理的思考力」「表現力」をどの程度身につけることができたか、あるいは何が身につかなかったかなど自分自身を振り返り、これからの学習に役立ててください。

第 3 部

[講演]

「学習成果」の設定と評価 —パフォーマンス評価を中心に— (2)

京都大学 高等教育研究開発推進センター 教授
松下 佳代 氏

1. ルーブリックの効用？

続きをお話ししていきたいと思います。今ご覧いただいたように、ルーブリックというのは、期待された、あるいは意図された学習成果を明確にする意味があるのです。最近、ルーブリックが広がりつつあるのですが、そのときに、ルーブリックの効用としてよく、ぶれないとか、時短になるということが挙げられています。ですが、これは本当かなと思うところがあります。

1) ぶれない？

ぶれないということに関しては、2つの意味があります。1つは評価者内信頼性、もう1つは評価者間信頼性ということです。例えば100枚近いレポートを採点していると、初めの方で採点したものと後の方で採点したものが、本当に同じ基準で評価したかなと心配になることが皆さんもおありかと思います。これについては、ルーブリックがあればぶれが小さくなる、つまり、評価者内信頼性が高まるということが言えると思います。しかし、もう一方の評価者間信頼性ということに関して言うと、複数の評価者間でぶれないというのは、なかなか大変なことです。先ほど見ていただいたように、レポート評価の場合、私たちもまだ成功しているとは言えません。実は、もう1つのPBL テュートリアルの方はかなり評価者間信頼性が高いのです。特に、実際に学生にロールプレイを実演してもらって、それを3人の評価者が評価するというものだと、高いものは相関係数が0.9ぐらいありますね。低いものでも0.7ぐらいです。ですので、単にルー

ブリックの出来だけではなくて、課題によって、パフォーマンスの解釈の可能性がより大きいものに関しては、どうしても評価者間信頼性が高くなりにくいと思います。先行研究もいろいろ見ましたが、レポート評価だと大体、一般化可能性係数で0.6ぐらいでした。

これが例えば1になるみたいなことは、実際にあり得ないわけですが、いろいろな先生の評価結果が多少異なるのはかまわないのではないかと、ある程度の許容範囲というところでおさまればいいのではないかと考えています。

2) 時短？

もう1つの時短ということに関しても、日本の教員の場合に、ルーブリックを使うことで本当に時短になるかどうかは、検討の余地があります。日本では、期末のレポートは出しっ放しで、教員からコメントを返さない場合が多いですよ。したがって、これまでは直観的に評価して点数だけ返していたのが、ルーブリックに沿って評価するということになると、かえって時間がかかるようになるかもしれません。それに対して、アメリカでは、これまで手書きでコメントを返してきたのが、ルーブリックを使うと、書かれてあるどのレベルにあてはまるかを示すだけでよくなるので、すごく時短になるわけです。あるいは、アメリカの教員というのはTAを使ってレポート評価をさせる人たちも多いわけですが、TAの人たちに、「このルーブリックで評価してね」と指示できるので、すごく効率的に評価できるわけです。ですので、そういうアメリカと日本のレポート評価の文化の違いというのを考えると、日本の場合には、時短というのが必ずしも売りにはならないだろうと私は思います。

2. 規準の設定と基準合わせが重要

私も実際に十字モデルを使って教えていまして、今日も午前中の授業でそれをやってきたのですが、ルーブリックで重要なのは、レベルをどう設定するかというところよりは、むしろどういう規準を設定するかということだと思っています。レポートを書くときにはこういう点に注意して書きなさいという指導をするときに、その規準、観点が有効だと思っています。このような指導の上での有効性だとか、あるいはほかの人に対して自分の評価基準を説明できること、そうい

ところが、教員からみたルーブリックのいい面なのだろうと思います。

複数の教員で評価をする場合や、あるいは、先ほどの「山口と世界」のようにルーブリックを共有する場合は、基準合わせがとても重要です。「基準合わせ」の原語は、キャリブレーション (calibration) です。例えば、時計を標準時に合わせる、時報に合わせるのをキャリブレーションといますが、それと同じように、ルーブリックについての理解を共有するというのがキャリブレーションなのです。単にルーブリックを読むだけで基準合わせをやるより、実際の学生の作品事例を、ルーブリックを使っていっしょに評価してみるというやり方で、基準合わせをやるとよいと思います。

とはいえ、先ほど小野先生が言われたように、採点者トレーニングでは、やはり限界があります。新潟大学の場合のレポート評価で、なぜそんなに信頼性が高く出ないかという原因を探ってみると、コピペに対する教員側の態度の違いというのがかなり関与しているようなのです。ですので、今、私たちの間では、一度、コピペ検出ソフトを使ってみて、どうなるか見てみようかと議論しています。

13. 〈誰が〉評価するのか

最後に、誰が評価をするのかということについてお話しします【スライド③-2】。



ルーブリックという形で評価基準を示すことで、学生自身による自己評価が可能になります。小野先生の資料【別紙資料】にも示されているように、実際、ルーブリックを使って学生に自己評価をやってもらっています。

この場合のメリットですけれども、1つは、学生の学習に水路づけを与える。どういうふうにやればいいレポートになるのかということを示すことができるということです。そして、それに向けて自分の学習をセルフ・レギュレーション（自己調整）していける評価主体として育てるという面があります。

一方で、デメリットとしては、評価に向けて学習するという、傾向と対策的な学習になってしまうという点がやはりあるのです。私がまた別に関わっている大学で、これはルーブリックではなくてチェックリストだったので、**「事前に見せられると、いい点を取るにはどうすればよいかばかり考えるようになるので、チェックリストは事前に見せないでください」**と言ってきた学生がいたのだそうです。それで、その先生はどうされたかということ、教員や上級生のいいパフォーマンスを見せて、そして自分のパフォーマンスもビデオで見せて、それらを比べながら、何が重要な点なのかを学生たちに考えさせてグループでチェックリストをつくらせた。そうして、できあがってきたチェックリストを見たら、教員側が準備していたものとかかなり似通っていたそうです。

そういうことで、ルーブリックという評価基準を示すことも重要ですが、それ以上に重要なのは、よいパフォーマンスに出会わせることではないかと思っています。私たちも、研究者として育っていく過程で良い論文に出会い、どのように書けばよいかということ、そこから学んでいきました。そういう体験をさせることが重要ではないかと思っています。

14. 〈学習としての評価〉という考え方

そのような意味で言うと、評価というのは単に学習の評価ではないのです【スライド③-3】。

「評価」と「学習」という2つの言葉をどのようにつなげるかについては、3通り考えられます。まずは学習の評価、従来から言われてきた assessment of learning です。例えば、成績の決定をするために、教師が評価主体として学習を評価するということです。

次に言われるようになってきたのが、総括的評価ではなくて形成的評価ということで、評価を授業改善とか学習改善に生かすということですね。これも教師が評価したものを学生にフィードバックして、学習改善に生かす。あるいは、教師が自分自身にフィードバックして、授業改善に生かすということだったわけです。したがって、これは学習のための評価、assessment for learning と言うことができます。

もう1つは、学習としての評価、assessment as learning です。学生が自分でも評価を行い、そして、自分の学習の自己調整のためにその評価を使っていくということですね。レポート評価ではそれほど感じられないかもしれませんが、卒業論文などはまさにそうだと思います。自分で書き直しをする、そして何回も人に見てもらって、そこから評価、フィードバックを受けて、また書き直していくこと自体が学習経験になるわけです。そういうふうな形で評価課題に取り組むことが、学習経験としての意味も持つ。そういうふうな評価をどれだけデザインできるか。評価それ自体が学びになるような評価をデザインするというのが今、大学教育に求められていることなのではないかと考えております。

以上で小野先生と松下の講演を終わらせていただきます。どうもありがとうございました。



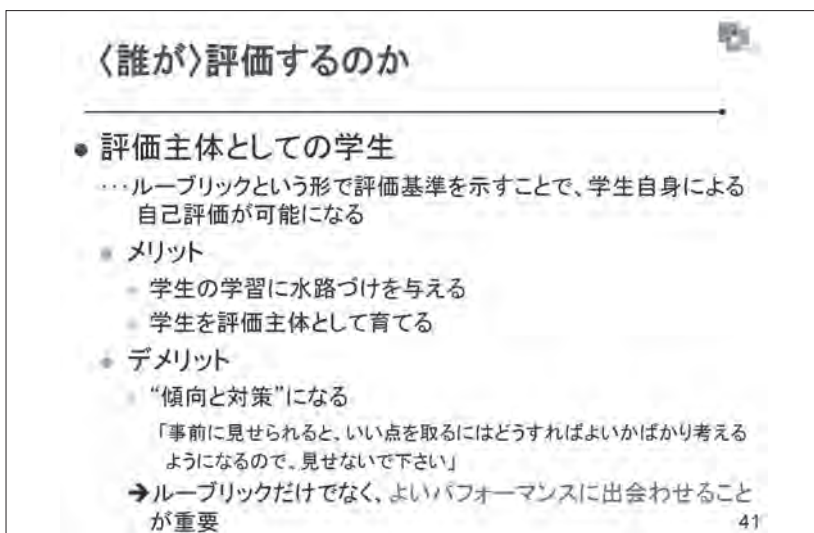
【スライド③-1】



学習としての評価

39

【スライド③-2】



〈誰が〉評価するのか

- 評価主体としての学生
 - …ルーブリックという形で評価基準を示すことで、学生自身による自己評価が可能になる
 - メリット
 - 学生の学習に水路づけを与える
 - 学生を評価主体として育てる
 - デメリット
 - “傾向と対策”になる
 - 「事前に見せられると、いい点を取るにはどうすればよいかばかり考えるようになるので、見せないで下さい」
 - ルーブリックだけでなく、よいパフォーマンスに出会わせることが重要

41

【スライド③- 3】

〈学習としての評価〉という考え方

- 3つの評価 (Earl, 2003; Alverno College Faculty, 1994)
 - 学習の評価 (Assessment of learning)
 - 成績の決定／教師が評価主体
 - 学習のための評価 (Assessment for learning)
 - 授業・学習改善／教師が評価主体
 - 学習としての評価 (Assessment as learning)
 - 学習の自己調整／学生も評価主体
 - +
 - 学習経験としての意味

評価それ自体が学びになるような評価をデザインする

42

【スライド③- 4】

ご清聴ありがとうございました



松下 佳代
matsushita.kayo.7r@kyoto-u.ac.jp

43

【スライド③- 5】

- 松下佳代 (2014b). 「共通教育における学習成果の直接評価—研究目的・研究計画・進捗状況について—」『大学教育学会誌』第36巻第1号, 62-66.
- 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・田口真奈・大山牧子 (2013). 「深い学習の評価ツールとしてのコンセプトマップの有効性—哲学系入門科目でのアクションリサーチを通じて—」『大学教育学会誌』第35巻第2号, 121-130.
- OECD (2009). *A Tuning-AHELO conceptual framework of expected/desired learning outcomes in engineering*. Tuning Association.
- 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』第36巻1号, 123-132.
- 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.

45

【スライド③- 6】

- 松下佳代 (2014b). 「共通教育における学習成果の直接評価—研究目的・研究計画・進捗状況について—」『大学教育学会誌』第36巻第1号, 62-66.
- 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・田口真奈・大山牧子 (2013). 「深い学習の評価ツールとしてのコンセプトマップの有効性—哲学系入門科目でのアクションリサーチを通じて—」『大学教育学会誌』第35巻第2号, 121-130.
- OECD (2009). *A Tuning-AHELO conceptual framework of expected/desired learning outcomes in engineering*. Tuning Association.
- 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』第36巻1号, 123-132.
- 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.

45

【スライド③-7】

- 佐藤望・横山 千晶・湯川 武・近藤 明彦 (2006). 『アカデミック・スキルズ—大学生のための知的技法入門—』慶應義塾大学出版会.
- 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.

○**司会** 松下先生と小野先生、ありがとうございました。松下先生からは、概念の詳細な説明、海外を含めた動向について、さらに、ループリックのメリット、デメリットについてお話しいただきました。小野先生からは、新潟大学での実践を通しての成功面と、ご苦労なさっている面の両面について、詳細な説明をしていただきました。今さらながら、本当に限られた時間で無理なお願いをしたなと思っております。それにも関わらず、お願いした内容に加えて、最新の研究動向まで踏まえてお話しいただき、先生方には、本当に120%の内容でご講演いただいたなという心境です。

今回は、学内外からたくさんの方がいらしていますので、この機会に、ぜひ質問をお願いします。これからの進行は、小澤先生にお願いしたいと思います。では、小澤先生、よろしくをお願いします。

○**小澤** お二人の先生には、大変素晴らしいご講演をいただきました。ありがとうございました。引き続きまして質疑応答に入らせていただきたいと思います。いかがでしょうか。



質問1

○**家城** 理学部の家城と申します。ありがとうございます。最後のところで少しお話がありました、ルーブリックにはいろいろメリット、デメリットがあるということ、学習としての評価という考え方ですね。それから、評価をする主体としては、学生というのも大きいということでした。そこでお聞きしたいのは、学生に対してこのルーブリックという基準をどういう形で提示するのがいいのかという点です。先ほどの小野先生のお話の中では、課題とともに配布しているということでした。ただ、配布しただけでは、先ほどデメリットと言われた、傾向と対策になってしまう可能性があります。うまく機能した場合には、学生がどのように学んでいったらいいかということ学ぶチャンスになる。ルーブリックをどういうふうに学生に見せるのがいいとお考えでしょうか。その辺で何かありましたら教えていただければと思います。

○**小野** 時間がなく十分にご説明できず申し訳ありませんでした。課題シートを学生に配布するときに、ルーブリックも一緒に配っていますというお話をしたのですが、少し追加すると、問題解決や論理的思考とは何かという講義をして、またルーブリックの記述語の意味するところを説明してから、ルーブリックを配っ



ています。まずは、少し補足させていただきます。

○**松下** 例えば、私は後期に教育学ゼミナールⅢ（学力・学校・社会）という全学共通科目を持っているのですが、今、「世紀末学力論争」というところをやっています。今日の授業では、荻谷剛彦さんの中央公論に載った論文一例の79年と97年の高校生のデータを使ってインセンティブ・ディバイドを指摘した論文ですが、あの論文の抜粋を学生に読ませて、何が問題で、どういう事実・データを使って、どういう結論を導いているか。彼が対立意見として想定していたのはどんなことで、そして、それに対してどういう論駁を行っているのかということ进行分析させました。先ほど、自分が書けるようになる前に、まずいいパフォーマンスを見せることが重要だと言いましたけれども、まずそういうふうな論文を読ませ、それを分析の素材として使うことで、このルーブリックの規準の中身を、十分学生に学ばせています。私は、規準というのは、そういうふうな期待される学習成果を分節化し明確化するという意味があると思っています。「問題解決」とか「論理的思考」とかいった一言であらわされるものが、実際にはどうということなのかということをかみ砕いたものが、ああいう規準になっていますので、それをまず授業の中でちゃんと教えることが必要です。ルーブリックそのものは、私は見せなくてもいいかなと思っているのですが、レポートを書かせる最終段階で、一応「こういう形で私は評価をしています」ということを示すために見せています。一方、規準自体は、授業のかなり早い段階で示して説明しています。

○**家城** ありがとうございます。今のお答えは、例えば、1つの科目の中でレポートを書かせるといった場合がそうなのかなと思うのですが、学士課程全体というか、卒業するまでにここまでという話もあったのですが、そういうものに関してはどういう形がいいでしょうか。

○**松下** 京大ではまだそういうことはできていません。先ほど、関西国際大学のコモンルーブリックをご紹介したのですが、あれは各科目で共通して身につけてほしいスキルを評価するときに使われているのですね。ですから、あくまでも各科目で使われている採点用ルーブリックと言えらると思います。ただし、先

日、関西国際大学でインタビューをさせていただいたときにうかがったのですが、もう一つ、長期的なルーブリックに近いものがまた別個あるそうです。半年に1回、学習アドバイザーと学生が1対1でそれを見ながら、ポートフォリオを使って、自分が今どういうふうな成長段階にあるかということの評価をするということでした。

○家城 そうすると、入学した時に学生にそれが提示されていて、学期ごとに何らかの評価をやっているということですね。ありがとうございます。

質問2

○佐藤 立教大学のやはり理学部なのですが、佐藤と申します。今の点に関係があるのですが、やはりこのルーブリックというのは、評価する先生にとっては便利なものだろうなというのはすぐ分かるのですが、学生に対して提示して学生がそれを通じて学んでいくという便があるというのはとても大事なことだなと思いました。

小野先生にお伺いしたいのは、バージョン1とバージョン2では、やはりだいぶ発展したなという感じがするのですね。学生から見ても、バージョン1のまま



ではどう書いていったらいいのか、やはり分かりにくいだろうと。そこでお聞きしたいのは、実際に採点されていて、ルーブリックを発展させたことによって効果が出ているという実感をお感じになったかということ、また、こういう努力がどのくらい学生にまで浸透しているのかということです。そこ次第では、やはりこういう評価型の実践だけではない授業改善が必要ということになりますよね。実際にレポートを何十枚も見られての実感のようなところを伺えたらと思います。

○小野 ご質問ありがとうございます。ルーブリックを使っていなかったころに比べて、うまく書いてくる学生が出てきたと言っていいでしょうか。正直な話、著しくよくなったということはないのですが、確かに、「学習のガイドとして、ルーブリックを意識して書いているな」ということが分かるレポートが出てきています。評価はそれなりに大変ですから、学生の成長を感じられなければ、教員は疲れるだけで、やがてはルーブリックをやめてしまうかもしれません。しかし逆に、教員が学生の成長を実感できれば、評価負担はあまり気にならないし、やる気も増すでしょうね。

「大学学習法」は、初年次前期に15週30回でやっているのですが、内容が



盛りだくさんで、ルーブリックで教員が評価した結果を学生にフィードバックする時間がとれていません。先ほど松下先生がアメリカと日本のレポート評価における文化の違いみたいなお話をされていましたが、ああいうことをやらなければいけないと分かっているのです。しかし、カリキュラムとか授業時間の関係でやれていないのが現状です。そこで、苦肉の策で、課題シートの一番下のところに、「自分で自己評価して、その結果を今後の学習に役立ててね」と書いてみました。教員の評価結果を学生にフィードバックしていないから、うまく書いてくる学生があまり多くならないのだろうと反省しており、今後、改善していこうと考えています。

質問3

○原田 法学部の原田でございます。お二方、本当に貴重なお話をありがとうございました。松下先生に質問させていただきます。私の教師としての理想は福山雅治でございます。確か京都大学でロケがたくさんあったと記憶しておりますが、あのテレビドラマで彼がよく言っていた言葉を今2つ思い出しているのですが、「実に面白い」と、もう1つは「さっぱり分からない」と。教師としての理想でもございますし、学生にもああってほしいと思うことが多々あるのです。た



だ、福山雅治演じる主人公は、to know や to do というのではなくて、not to know というか、まず「分からない」と言ってしまう。これが学生の場合、ずっと「分からない」のまま終わってしまって、次の実践といいますか、能力が発揮された現象として捉えられる段階にまで行かない場合も多いように思うわけでありませう。そうした、能力というよりは意欲というのか、あるいは動機というのか、学ぶに当たってはそちらも非常に大事で、私としてはそれを学生にどうやって植え付けたいののだろうかと思うところもあるわけませう。そうした問題とループリックというのは、どのように関連づけたいのかということをお聞かせいただけるとありがたいのですが。

○**松下** 非常に本質的で、かつちょっとお答えしにくいご質問をいただいたと思います。「実に面白い」、「さっぱり分からない」と言いながら、福山雅治は、最終的には問題を解決していますよね。「実に面白い」、「さっぱり分からない」と言って、さっぱり分からないままだったら、何のすばらしさもないと思うのです。「実に面白い」、「さっぱり分からない」で終わってしまわないで、分からないからこそ探究し続けて、最後は一応、分かった段階まで行っているから、あれは番組として成り立っているわけだし、あの先生のすばらしさがほかの学生たちにも分か



るわけで、そういう意味では、福山雅治はすごく to do していると思うのですね。やはり何らかの形であるすばらしさを観察可能な形にしているものがあると思うのです。ただ、これが研究であれば、「さっぱり分からない」と言いながら探究し続けても、実際、答えにたどり着けずに終わってしまう場合もあると思います。フェルマーの定理とか、長い時間をかけて挑戦しても、最終的に証明できずに終わってしまった人はいっぱいいるわけですがそれでも、でも、その途中で何らかの成果を生んでいる場合もあるわけです。私は、さっぱり分からないけれども、分かろうとし続けることが重要なのではないかと考えています。

質問4

○**学外者** **と申します。本日は貴重なご講演をありがとうございました。1点、松下先生へのご質問になります。時間の関係であえて割愛をされていたかと思うのですが、標準テストへの批判の内容について、パフォーマンス評価との対比として教えていただけると非常にありがたいなと思っております。よろしく願いいたします。



○**松下** 標準テストへの批判というのは、これはあくまでもアメリカの標準テストの話です。CLA、Collegiate Learning Assessment についてはご存じかと思いますが、OECDのAHELO (Assessment of Higher Education Learning Outcomes) でもジェネリック・スキルの評価で使われたものです。私は今回のAHELOのジェネリック・スキルの問題を、ちょっと分析してみましたけれども、問題自体はとてもよくできていると思います。でも、結局、AHELOのジェネリック・スキルの評価というのは、あまりにアメリカ的過ぎる。文化的に偏り過ぎていているということで、国際調査という点から見るとちょっと問題があるということになったようですね。

標準テストへの批判ということで、VALUEプロジェクトのときになされていた批判をご紹介しますと、CLAというのはフルでやると1年生、2年生、4年生で実施します。その大学の学生全員に対してやるわけではなく、抽出でやるのです。ですので、教員側からすると、抽出された学生についての標準テストの結果では、大学全体のことは分かるかもしれないけれども、自分の担当した学生がどうなのかということとは分からない。それから、ジェネリック・スキルのテストだと、結局、自分が普段教えていることがどれだけ学習成果としてあらわれているかも分かりません。こういった点が、VALUEプロジェクトで出されていた一番の批判点です。ですので、教員から見て、もっと自分たちが教えたことがどれだけ学生に学習成果として身につけているのかを直接見られるような評価方法を編み出す必要があるということで作られたのが、VALUEルーブリックだったのでね。

今日はVALUEプロジェクトについてはルーブリックの話だけでしたがけれども、そのときに評価対象になったのは、eポートフォリオに収められたいろいろな学生の作品です。日本とアメリカでは、またちょっとeポートフォリオの普及や使い方が違うと思いますけれども、標準テストへの批判として出されたのは、今申し上げたようなことです。

質問5

○**栗田** 理学部の栗田です。理学部ばかりですみません。私は物理学科の学生を対象とした導入科目で、スモールグループ・ディスカッションを導入しているのですけれども、そういうグループ・ディスカッションの中でも、恐らくこのルー

ブリックというものを設定することは可能だと思うのですね。しかし、十何グループある中で、私は全体を把握することはできていないわけです。そういう教員が全体を把握できない状況で、このルーブリックというものを使うことは可能なのでしょうか。よろしくお願いします。

○**松下** 1つは、ピア評価と自己評価のためにルーブリックを使うことはできると思います。今、ピア評価も結構凝ったものいろいろ出ていまして、今日はまだ詳しくお話しする時間はありませんが、例えば Calibrated Peer Review といわれるものがあります。これは、まず学生にどの程度アセスメントの能力があるかを見た上で、アセスメントの能力が高い学生の評価は高く重みづけをし、低い学生は低く重みづけをすることでピア評価の結果に妥当性を持たせようとするやり方です。もしご興味がおありでしたら、Calibrated Peer Review で検索していただければすぐヒットすると思いますので、ご覧いただければと思います。

○**栗田** ありがとうございます。その学生のピアレビューで成績はつけられるのですか。



○**松下** そこが難しいところで、それこそが Calibrated Peer Review の開発に至った背景ですね。学生の自己評価、ピア評価において、それぞれの学生にどのくらいの評価能力があるのか。それによって評価が変わってくると思うのですね。ですので、Calibrated Peer Review の場合は、例えば、学生にレポートを評価させる前に、まず教員が評価した学生のレポートを3本、学生に評価させてみます。とても出来がよいものと、真ん中のものと、あまり出来がよくないものの3本です。学生に評価させてみて、ちゃんと評価できれば、その学生は評価能力が高いということで、その学生の評価した結果は高く重み付けをするようにしているのです。そのやり方がいいかどうかは分かりませんが、まず学生の評価能力を評価した上で、今度は学生どうして評価させて、そのピア評価の結果を成績に反映させるというやり方がとられています。それについては、今、私たちのゼミでもちょっと検討してみようかと言っているところで、メリット、デメリットについてまだ十分検討できていないので、今日はこの程度しかお答えできないのですけれども。

○**栗田** ありがとうございます。

○**小澤** まだまだご質問があるかと思いますが、そろそろ時間になりましたので、以上をもちまして質疑応答、ディスカッションの時間を終わらせていただきたいと思います。松下先生、小野先生、ありがとうございます。

閉会挨拶

○原田 センター長の原田でございます。本当にお二方、今日は大変貴重なお話を頂戴いたしまして、ありがとうございました。私どもは、ちょうどあと1年半後でございますが、10学部を横断するような新しい学士課程のプログラムをスタートさせるべく準備を進めているところでございます。また、スーパーグローバル大学創成支援事業でも、本学は採択をされましたが、その中でもやはりルーブリックの導入ということについて記載をいたしております。そうした状況もございまして、これからどうやって私どもがルーブリックに、あるいは教育の質改善に向き合ったらよいのかということについて、毎日毎日、疑問に直面しているのでありますが、本日は非常に参考になるお話を伺うことができました。今日のお話をぜひとも全学で共有いたしまして、2016年に向けたカリキュラム改革、さらには今後10年間のスーパーグローバル大学の創成支援事業をしっかりと進めて参りたいと存じます。お二方には、これからもご指導くださいますよう、どうぞよろしくお願いいたします。



大学教育開発・支援センター センター長、副総長、法学部教授

原田 久

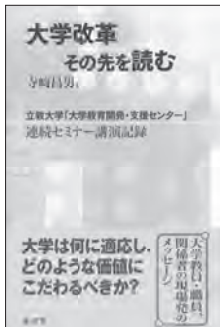


●大学教育開発研究シリーズ バックナンバー

- NO.1** 外から見た立教大学
—ミッションと社会的要請— (2006年)
- NO.2** 「学生による授業評価アンケート」にもとづくRIKKYO授業ハンドブック
—学生の積極的な学習を励ますために— (2006年)
- NO.3** 変化する高校生と大学への期待
—高校から見た立教大学— (2007年)
- NO.4** わが大学・わが学部の教育改革を語る
—学生の学ぶ力、選ぶ力とカリキュラム— (2007年)
- NO.5** 立教大学の初年次教育とその展開
—〈勉強〉から〈課題探求型学習〉への道— (2007年)
- NO.6** 学生が見た立教大学の初年次教育
—今後の充実に向けて— (2008年)
- NO.7** 立教大学の今後と中教審の審議
—学士課程教育の再検討と将来を考える— (2009年)
- NO.8** バージニア工科大学視察報告
—米国における先進的な教育改革の事例に学ぶ— (2009年)
- NO.9** 立教大学における学習支援と図書館 (2009年)
- NO.10** 立教大学におけるアドミッション・ポリシー (2010年)
- NO.11** アクティブな学びをデザインする
—4つの授業をめぐる対話— (2010年)
- NO.12** グローバル化に対応する大学教育の在り方
—東アジアの高等教育における質改善への取組に学ぶ— (2010年)
- NO.13** 大学生の社会的・職業的自立に向けた教養教育の在り方 (2011年)
- NO.14** アクティブな学びをデザインする vol.2
—学生の気づきを促す3つの対話— (2011年)

- NO.15** 学位取得へ導く大学院教育のあり方
—博士後期課程を中心として— (2012年)
- NO.16** 日本の大学に求められている国際通用力とは (2012年)
- NO.17** 学びが高まる学習環境とは
—ハード、ソフト、コミュニティー (2013年)
- NO.18** 大学院研究指導への誘い^{いざな}
—海外マニュアルの紹介— (2013年)
- NO.19** 「読む」学生が育つ大学教育を求めて
—若者の読書実態と授業実践を始点として— (2013年)
- NO.20** アクティブな学びをデザインする vol.3
—2学部における初年次演習科目の実践から— (2014年)
- NO.21** アクティブな学びをデザインする vol.4
—学びの転換を促す「導入期」演習科目— (2014年)

●連続セミナー講演記録



寺崎昌男『大学改革 その先を読む』(2007年)
東信堂 ¥1,300

編 集

御手洗 明佳 (大学教育開発・支援センター 学術調査員、2014年12月まで)

谷村 英洋 (大学教育開発・支援センター 学術調査員)

大学教育開発研究シリーズ NO.22

「学習成果」の設定と評価

—アカデミック・スキルの育成を手がかりに—

2015年3月発行

発 行

立教大学 大学教育開発・支援センター

〒171-8501 東京都豊島区西池袋3-34-1

TEL : 03-3985-4624 FAX : 03-3985-4615

<http://www.rikkyo.ac.jp/aboutus/philosophy/activism/CDSHE/>

e-mail : cdshe@grp.rikkyo.ne.jp

制 作

株式会社アクセスリード

〒150-0002 東京都渋谷区渋谷2-15-1 渋谷クロスタワー 24階

TEL : 03-5774-2330 FAX : 03-5774-2339



RIKKYO UNIVERSITY