

米国の中規模州立大学における学習成果の診断 (Assessment)

本田 寛輔

Director of Assessment

Office of the Provost

University of Maine at Augusta

米国の中規模州立大学における 学習成果の診断 (Assessment)

大学評価担当者集会
2015年11月17日

本田 寛輔

Director of Assessment

Office of the Provost

University of Maine at Augusta

日本でも何回か講演を引き受けてきましたが、帰ってくると、いつも浦島太郎状態です。以前は「日本で事務職をしていました」と最初に伝えれば、参加者の方々からは「あの人は大丈夫かな」と思っていただけたのかもしれませんが……。アメリカに10年ぐらい居ると、日本にたまに帰ってくると新幹線のチケットの買い方が分からなくて、恥ずかし

い思いをしています。

日本での講演を引き受ける際に、幹事の浅野先生や、寫田先生に「どういお話をしたら日本の方に有用なんですか」とお話をいつも伺っています。ところが、当日の講演で話の内容が参加者の期待とうまくかみ合っていないというのは、皆さんの顔を見れば分かるわけです。大講堂でみんな目が点になっている中、以前は冷や汗をかきながら講演してきました。

そこで、今回は普段のような講演ではなく、少人数で対話型の会にしたいと寫田先生にお願いしたところ、「もう本田節でやっていいよ」と言っていただきました。そこで、今日は自分が実際やってるデータ発表や、企画書はどんなの書いて、学内の調整をしながらやってるのかという、具体例をどんどんお見せしながらやっていこうと思いますので、写真撮影はご遠慮願います。正直、学内でも自分の上司の教学担当副学長には見せても、学科科長レベルにはあまり見せないようなデータもあるので。

はじめに

- **講師の自己紹介**
 - 全入、社会人、遠隔教育の大学
 - ホテルのレストランから町の食堂へ転職
- **発表の目標**
 - Assessmentの概論と手法
 - Assessmentの結果の活用
 - Assessmentの活動実態

まず最初にお断りさせていただきたいのが、講演のご案内の時点から、演題を『米国の中規模州立大学における学習成果の診断』と変えさせていただきました。

次に自己紹介ですが、私も日本で事務職をやっていたときは、アメリカに行って成功事例を日本に持ち帰ろうと考えていました。青い鳥を探して行って見たのですが、実際に

働いて、もう7、8年になりますが、現実はどうでもない…。日本から見たときに、アメリカの大学と言うと、一緒くたになってしまうときがあるようですが、皆さんもご存知のように私立のハーバードがあれば、地域の州立大学、無名の小規模私立もあります。様々な種類の大学が存立しているのは日本も同じでしょう。なので、今回の私の話は中規模の州立大学に行われている学習成果の診断が中心となります。正直、大規模有名大学は入学の選抜度が高いことや、教員が研究志向なこともあり、地域の中小の大学よりは学習成果の活動が遅れている節があります。

また、米語ではアセスメントと言いますが、日本語だと「環境アセスメント」の印象が強いのかと思います。私なりには「診断」と意識しています。企業診断のような意味合いです。

話を私の現在と前の勤務校について移します。アメリカだと各州に大体一つぐらいはありますが、社会人で仕事をしている人を対象にオンライン授業やサテライトの学習センターを通じて遠隔教育を提供する大学です。そうした大学というのは、概ね全入に近いところがあります。というのも、18歳、22歳の間に学業を終えることができなかつた人達の為の大学ですから。この辺を今日の話の背景としてご承知おきいただければと思います。

では、本田の話は無名校で全入の大学だから、あまり参考にならないのかということ、実はそうでもないと思います。何が興味深いのかということ、社会人対象の全入で遠隔教育という状況では、退学者が多く出ます。仕方が無いのは、学生は平均33歳の社会人で、自分の仕事だけでなく、子供や親の面倒を見る必要があり、予期せぬ事が学業に支障を来します。例えば、ある授業に100人登録したら、20人が学期末までにいなくなるとか。その分、教員はもちろん、執行部もあの手この手で学生の支援に必死です。更には、全入の大学は他の大学からも多少厳しい目で見られるわけです。例えば、勤務校の学生が1つ格上の大学に編入したり、卒業後に大学院へ進む際に「この大学は教育をしっかりとってるのか？」というふうに懐疑的に捉えられがちなのは否めません。こうした逆境に立たされているからこそ、自分の勤務校が他所の大学よりも頑張っていることは結構あるのかなと思います。これは半分ひがみに聞こえてしまうかもしれないのですが、知り合いなんかを通じて有名大学のIRやアセスメントの部署に働いてる人のお話を聞くと、「あまりにも事務処理業務で、自分たちのデータ分析が何の意味決定に使われているか分からない」そ

うです。「結構、つまらないよ」といった節の話を聞くと、自分は今の勤務校で経験を積んでおくのはいいのかなと思っています。

2つ目の項目で、「ホテルのレストランから町の食堂へ」というのはたとえ話です。私のIRとしてのキャリアは大学院生の事務補助員としてニューヨーク州立大学アルバニー校で始めました。ここは学生数が17,000人くらいの大学で、IR部署は7、8人の比較的大規模な所帯でそれぞれの業務は専門分化され、一つの蝸壺になっていました。その後、常勤で同じくニューヨーク州立大学の遠隔教育を中心とするエンパイア・ステイト・カレッジでも同じようなIE部署で、自分の任される業務範囲は狭く、隣と同僚が何をやっているのか十分に理解できていませんでした。

これでは幅が広がらないと考え、メイン州の5,000人くらいの学生数で、3人のIR部署に移りました。「ホテルのレストランから町の食堂へ」というのは、去年の大学評価担当者集会の畠田さんの例え話にも触発されてなんですが、ホテルのレストランというのは、ニューヨーク州でのIR部署のように規模が大きく専門分化されています。そこでは、「俺、キッチンが中心で、接客したことないよ」というのが普通にありえます。ところが、今やメイン州の町の食堂に来てしまった感じです。お客さんは食べたい物が分からないのに、「とりあえず何か適当にやってよ」と注文して、こっちが工夫して食事を提供すると。「これ、何か違うなあ」なんて平気で言います。そういうお話も各事例の中でお話しできればと思います。

次は発表の目標です。これは私なりの反省が含まれています。初めて小湊先生に九大にお呼びいただいたときは、すごい頑張ったつもりです。私なりに日本の大学改革にいくらかでも貢献できればと気張って、最初の概念的な説明が多くなりました。日本の方々には二つ目の「結果の活用」を知りたいのは分かってるのですが、結果の活用の全体像が見えているのは、ディレクターや副学長など、上のほうなんです。私の場合、5年前はそういう管理職に就いてなかったもので、データは学科長に提供してるけど、副学長クラス的意思決定については限られた情報しか得られませんでした。というのも、規模の大きな大学は管理職の階層がそれだけ厚くなり、データ以外の要素が意思決定を左右するケースが多いからです。

今の大学に移って管理職になり、だいぶ視線が変わったなという気がしています。もう今は学部長だとか、プロボストという、アメリカの大学では複数の副学長の中でも筆頭の教学担当に付いています。19人の学科長の他に、様々なプロジェクトごとに各部署のディレクターなども取りまとめて仕事を進めるので、管理職としての自覚というのか、単に上職者を支援するだけでなく、場合によっては集団を引率する度量が試されるようになりました。

今日のお話は綺麗な成功例ではなく、どちらかというと3つ目の項目の「活動実態」です。手の込んだデータ分析したけど、あまり認知されなかったとか、小さいデータのまとめをやったら意外に喜ばれたとか、そういうお話ができたかなと思っています。よろしいでしょうか。

学習成果が求められる背景

- 学費の上昇
- 成績のインフレ
- 就職率の低下
- 政府の規制

2

追記：当日はこのスライドの説明を飛ばしました。米国の総論的な背景はスライドをご覧ください。推察できるかと思います。

学習の成果って何ですか？

コミュニケーション

就職率

IT技能

専門知識

就職先

批判的思考

大学院進学

文章作成

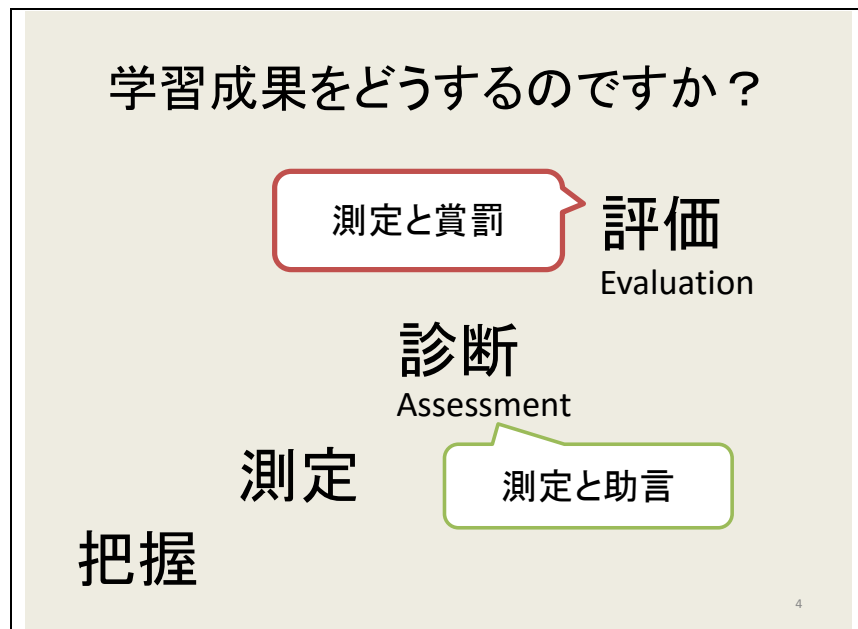
3

このスライドと次のスライドの内容を十分に理解し、適切な分析手法を提示できれば、今日の講演はこれ以上聞かなくても大丈夫なぐらいの根本的な問いです。

さて、「学習成果って何ですか？」最初に学習成果を英語にすると **Learning Outcomes** です。日本で働いた経験もあるので、学習成果って何かというと、スライドの右側にあるよ

うな就職率や就職先が真っ先に浮かんでくると思います。片や、スライドの左の方には技能や知識の習得が並べられています。アメリカでは、どちらかというと、左側の批判的思考だとか、文章作成だとか、そういう学生が実演可能な能力で、かつ、測定ができるものを学習成果と呼ぶようにしています。だからといって、アメリカが就職率や就職先のデータを取っていないわけではありませんが、日本ほどデータが取れてるわけではありません。日本だと就職率は新聞などで大きな話題になりますが、アメリカは実際いろんな統計ありますけど、40 何パーセントしか就職できていないようです。そのような状況で就職率を学習成果に捉えよと言われても、正直難しいです。また、例えば大学で批判的手法を教えたから、就職率が上がるのか、そこに 1 対 1 の因果関

係があるかといえば、そうでもない。大学教員は学術的な技能を教えれば、それでいいという割り切りがアメリカにはあるのだと思います。この辺の話は大丈夫ですか？成果といっても、いろいろあるからという話です。なので、多様な成果をどう定義して、測定し、改善に持っていくのかは、もう職人芸の域なわけです。



次に、「学習成果をどうするんですか」という話です。まず最初に、アセスメントという片仮名を持ち出すことについて説明させてください。アメリカだとエバリュエーションとアセスメントは違うんですね。日本だと、環境アセスメントという言い方が普及していると思います。ところが、アメリカの学習成果の文献でよく紹介されるのが、アセスメン

トの語源は「隣に居て、助言を与える」という意味があるそうです。例えば、私が今講演しています。隣に司会の山本さんがいらして、私の講演を傍で見てくれているわけです。そして、講演会後に「本田さん、あその話はあまりよく伝わってなかったですよ」だとか、「あの部分は具体例を用いて説明していただけたら、参加者の理解が進んだかもしれません」とか、私の講演の技術の上がるようにフィードバックをくれるというのが、アメリカの高等教育界におけるアセスメントの意味なのです。片や、日本でよく使われる評価ですが、英語での本来の意味は、「何らかの判断があり、その結果が、査定や昇格、財政配分に使われるなど、何らかのひも付けがなされる」ことです。日本とアメリカを比較する際には、基本的な考え方の違いを抑えてもらい、私の業務では、アセスメントを評価と取り違えられないよう配慮しています。

皆さん、まだピンと来ていないようなので、もう一つ踏み込んだ話をします。評価って非常に難しいんですよ。先程のアセスメントの例に戻ると、私が山本さんから講演に関するフィードバックをいただいたのを、自分で内容を精査し、改善に取り入れるわけです。アセスメントはそういう柔らかい感じのものなのです。ところが、私の講演を山本さんが評価するとなったら、「え、ちょっと待ってください。それで私の今日の謝礼の金額、変わるんですか？」となります。更に、「じゃあ、ちょっと判断基準を紙に書いてくださいよ」とお願いします。例えば、日本語が下手だとか、片仮名が上手に訳せていないだとか、評価項目を提示してくださいという話に発展します。まだありますよ。次に、「何を判断基準にしているんですか？」という、いわゆるベンチマークの話です。例えば、高等教育学会に出てる先生方と僕を比較して、どうなのか。それとも評価担当者集会の他の人の発表と比べてみるのか。測定方法は参加者のアンケートのみですか？それ

とも、幹事の方々の参与観察でルーブリックによる評価でしょうか。そこまで頑張って詰めていないと、評価ってできないですよ。

日本は政策の流れで言うと、質保証の制度として 90 年代に自己点検から導入し、各大学の自助努力に任せていたのが、それでは改善に結びつかないという批判が出ました。そして 2002 年の第三者評価の答申が求められたわけですね。ですから、「評価」という言葉の意味と文脈がアメリカとは違い、文科省の政策的な用語で使われ始めました。それは理解できても、アメリカにいったん渡った者から見ると、エバリュエーションは物凄く難しい仕事なのです。例えば、いくら私がアセスメントの専門家でも、社会学科の教授会に出て、「あなた方の授業は批判的思考の教え方になってないよ」とは言わないわけです。そうではなく、「批判的思考を教えるのであれば、このような概念整理や測定方法、教え方がありますよ。どのようにしたら授業の改善が進むでしょうか」とやんわりしたフィードバックをあげるのが仕事です。

	評価 (Evaluation = Measurement + Final Judgment)	診断 (Assessment = Measurement + Improvement)
プログラムの次元	GPA: 大学が奨学金、卒業、大学院進学などの判断の為に、科目の成績を平均化	カリキュラムの改善: 教員集団がカリキュラム改善の為に、学生の達成度をプログラムの学習成果に基づいて測定
科目の次元	成績: 教員が、学生の達成度を判断する為に、科目の教育目標に基づいて測定	教授法の改善: 個々の教員が教授法改善の為に、学生の達成度を科目の学習成果に基づいて測定

今日はいつもの調子で総論で 1 時間ぐらいつぶして、参加者ががっかりされないよう、各論からいきます。ただ、各論からいくと、どういう問題が起きるかですね。

例え話ですが、壁や窓の話が先で、土台になる話が後になります。具体例を出してご説明すると、私がこんな学生調査しました、このテストを導入しましたとか説明します。

そうすると、ここの窓ガラスは大きいですね。この壁の色は綺麗ですねという部分的なものが見えても、土台が見えていないと、何故ここに窓があり、壁の色を選んだのかという全体像が分からないかもしれません。それが各論から始めるときの、怖さですね。取りあえず窓の大きさと壁の色の話から始めてみて、最終的に土台、この建物はこういう形だったから、この窓とこの壁がこういう形になってるんですよというふうに最終的にまとめていければと思っています。よろしいでしょうか。

学習成果に関する業務

- 教養教育 (Gen Ed) の診断
- 学科別自己点検と外部審査の支援
- 学生による授業評価の実施
- 学習実態調査の分析
- EM関連の分析
- 他部署のアンケート実施の支援

6

追記：当日はこのスライドの説明を飛ばしました。

事例の一覧

- GPAの活用状況
- カリキュラム・マップ
- 学習実態調査
- 学生による授業評価
- ルーブリック
- 情報リテラシー共通試験
- CEQプロジェクト
- 戦略計画

8

これらが、今回ご紹介しようと思っている事例です。GPA や学生調査、カリキュラムマップ、ルーブリックとありますが、ちょっと挙手をいただければと思います。GPAを導入している大学は？半分くらいですか。GPA で卒業判定の数値が決まっている大学はありますか？例えば GPA2.5 以上でないと卒業させないだとか、進級させないよという所

はありますか。ない…。ということは、日本の GPA は単に成績を数値に置き換えてるだけですかね。

小湊：日本でも卒業要件に GPA を取り入れてる大学はあるけど、少ない。

本田：まだそんな状況だったのですか。私の GPA 活用の話ですが、勤務校で学習コーチというプログラムを試行中で、それを入れたことによって退学者がどれくらい減ったかというデータ分析に GPA を使ってます。GPA というよりは、成績分布のデータと言った方が正確ですが。要するに、学習コーチの導入前と導入後で、科目別の学生の成績分布を見て、A を取った学生が何十パーセント、B を取った学生が何十パーセント、C が何パーセントと変化があったのかをまとめました。

では、学習実態調査に関してですが、何らかの学生調査は実施しているという大学は

どれくらいありますか？それが改善につながった事例を、ご自身の役割から知っているって
いう方は？

私が今回ご紹介できるのは、学生調査でも NSSE と呼ばれる全米の学習実態調査を 2007
年、2010 年、2013 年で 3 回取ってあるので、そのトレンドを分析しました。分析した後
に、教員にどうフォローアップしたのか、学習経験の中で弱みである項目を全学目標とし
て来年度設定するという作業の途中段階なので、その話もできます。

あとは、情報リテラシー共通試験というのは、アメリカでよくある業者のテストです。
日本ではまだ少ないと思うので、このお話をして参考になるか分からないですけど。例
えば、図書館員と教員の協働の形をどのようにつくっていくのかという話もできます。

あとはカリキュラム・マップというのは、皆さん今やろうとしている、または始めてい
るところですか。この会の幹事の方から聞くと、「うちやってるよ」というのが多かったの
ですが、カリキュラム・マップっていう名称をご存じの方？ やって見たら凄くいいこと
あったという方？

小湊：若干。

浅野：実際にやるのは、大変です。

本田：うちの大学でも、経営学科は半年かかったと言いましたね。私の上司は、「10 分でできる
だろう」と言いますが、無理です。やっつけで単に穴埋めするなら、教授会で法令遵守的に
済ませることはできますが、真剣に議論したらやはり時間が掛かります。

ループリックの話ですが、これは私も 2 年前かな。大学評価・学位授与機構で既に講演
したので、ループリックが何かということと、どういうふうな測定をするのかっていうの
は、ちょっとご紹介するのみに留めさせていただこうと思います。というのも、実際の測
定をどうやるのか完全に説明するには、もう 1 つの講演では網羅できませんから。恐らく、
私がコンサルタントみたいに入って、月 1 回か 2 回のインターネット会議を持ちながら進
めたとしても半年から 1 年は掛かるでしょう。なので、今日の講演ではあまり込み入った
プロセスの話はないと思います。

さて、皆さん大好きな学生の授業評価です。これも、本当に改善に使えるのかが問題で
すが、苦笑いが結構見えるので、多分日米そんなに差はないんだと思います。

では、他には戦略計画だとか、全学の学習成果の活動はおまけというか、付録の資料で
す。話のネタとしては面白いと思いますが、「ああ、アメリカはそうなんですか」で終わっ
ちゃうかもしれないので、後半に万が一時間があるようならご紹介します。

順番ですが、自分が話したいネタと、皆さんがお聞きしたいネタに多少のズレがあるの
かなと思うので、カリキュラムマップから始めるのが妥当なのかなと思いますが、いかが
でしょうか。それで、GPA に移り、学生アンケートという流れでどうですか。

まだ前置きが続きます。そうだ。寫田さんに言ってもらったことで、冗談を言おうと思
ってたんだ。「本田節でいいよ」って言われたときに、幾つかあって、くどい、説教くさい、
冗談がつまらないときがあります。苦しくなった方は、トイレにでも行くようなふりして、
抜け出すのは OK です。

学習成果の測定道具

日本で米国の手法を導入しても、
上滑りしていませんか？

道具は使い方次第で効果が変わります。

「日本で米国の手法を導入しても、上滑りしませんか」ということですね。実は、日本だけじゃないんですよ。そういう比較研究とかも大学院の頃にやっていたので。例えば、歴史的にはドイツがイギリスの真似をして上滑りだとか、他の国の事例もあります。日本は特にアメリカへの関心が高いですよ。だから私も留学して移住しているのですが・・・。

さて、「上滑りってどういうこと？」いろんな次元であると思うんです。政策の謳い方だとか、政策の施行の段階できちんと情報提供できてないだとか、専門的な人材の養成ができてないとか、そういうのは高等教育学会系の話なので、ここではやめときますが。

私なりの最終的な結論としては、道具は使い方が変わると思うんです。私もアメリカで働いて、専門職や適格認定団体の研修会に行って、じゃあ、ルーブリックってこうやるんですよ、カリキュラム・マップはこうやるんですよと学ぶ機会はいくらでもあります。ところが、実際に学内で始めてみると上手く行かないときもあるし、始めたばかりの頃は「何だよ、これ、使えねえじゃん」って思ったこともありました。それでも、何年か頑張ってみると、要するに道具は使い方次第なんだなって気がきました。なので、皆さんも道具を見つけてすぐに使えるかという、そうでもないのでしょうか。2、3年は掛かるだろうと覚悟しておくのが現実的かもしれません。

どこまで出来ているのか？

答えは主観的¹であり、

且つ相対的²である。

1. 理想、期待値など
2. 比較対象、時系列など

10

もう一つ、事例に入る前に、よく日本で受ける質問で、「アメリカはどこまで出来ているのですか」と聞かれます。私が思うに、答えは主観的であって、相対的である。例えば主観的というのは、今日、私が講師として話しているので、大学の体裁を守るために「オーガスタ校はよくできています」というふうな私の主観で話すかもしれないし、逆に私が日本語で話してい

て、どんな暴露話や愚痴を言ってもバレないから、「あまり出来ていません」という主観で話すかもしれない。同じ問いを、よその何とか学科長に聞いたら、またその人の主観で判断されてしまうわけです。例えばある老教授、アセスメント委員会の委員長に聞いたら、「うちの大学は全然駄目」って言いますね。「俺は、ここの大学に20年勤めてて、アセスメントなんて全然できてないよ」と学内のあらゆる人たちをこき下ろします（これは、個人の性格の問題なのかもしれません）。そこで、彼に「何故そう思われるのですか」と聞くと、「だって、アメリカのどどこど大学と比べたら、うち、全然できてないじゃないか」と言うのです。「でも、先生、そうしたグッド・プラクティスと呼ばれる大学は何千校のうちの数校ですよ」と私なりに切り返しますが。

「且つ相対的」というのは、どの大学と比べているのかっていうのと、あとは自分の大学の中で時系列に捉えること。5年前、10年前と比べて今がどうか。ちょっと自慢話になって申し訳ないのですが、私は現在の勤務校に転職して2年で、日本でいう認証評価を通るぐらいまで学習成果の業務を充実させたのだから、十分にできてると思えるかと思います。もう一つ最近良かったのは、連邦政府の助成金の申請書を書くときに、評価の設計を必ず入れないと駄目なんですね。政府のお金もらったら、どのようなプログラムを導入して、どういうふうに査定したり、評価するのかという計画を記載する部分を自分は担当して、その項目は満点をいただきました。自分がやってる内容は、連邦政府の審査員の目に適うぐらいのことはできてるのかなという自己評価ができました。

カリキュラム・マップ

- 誰が、何を、どの程度、教えているか
<http://assessment.uconn.edu/primer/mapping1.html>
http://www.ynu.ac.jp/education/plan/curriculum_map.html
(日本語の参照例)
- 学内の文脈(専門課程と一般教育課程)
- 政治学科目の例
- 学習成果の診断への準備運動

11

さて、GPA は飛ばして、カリキュラム・マップからいきましょう。ごめんなさい。配布した印刷の資料だと飛ばし飛ばしになってしまいますが。まず、「誰が何をどの程度教えてるか」という話ですが、結局カリキュラム・マップって何なのかというと、大体技能にはいろいろな種類があるわけですね。文章作成能力、IT 能力、あとは批判的思考なんてい

うのもあります。そういった技能があり、どの授業で教えてるのか。教えてる内容の次元を説明するのが一般的なやり方です。例えば、導入的に教えています。Intermediate は訳しにくいですが、中級程度に教えています。そして、Advanced は上級と簡易的に訳しましょうか。

これを、教員が教授会で集まって、例えば「私の授業では数値の分析に関しては導入程度です」など、情報を共有して誰がどの科目で何を教えているかを明らかにしていきます。日本でも、確か横浜国立大と金沢工業大がインターネットで公表していた記憶があります。今回の印刷してあるパワーポイントは皆さんにも電子版をお送りするので、そのリンクを見ていただくと正確な情報が取れます。

私の大学はちょっと違うやり方でカリキュラム・マップを作成しています。まず、初級、中級、上級というふうに簡単に訳させていただきますけど、アメリカは科目の番付けがあり、100 番台は1年生、200 番台は2年生、300 番台は3年、400 番台は4年と大まかに分かれてるので、そこをわざわざ初級なのか、中級なのか、上級なのかということは明記していません。それよりも、各科目の中で、どのような授業内容が展開され、どのように測定しているかを示すような作りになっています。多分、この例はアメリカでもあまり例がないと思います。だから、「凄い」とか、「良い」とか、「進んでいる」と言うつもりはありませんが、私は勤務校のやり方が各教員が何をやってるのがよく分かるので、実践的だと思います。例えば、「批判的思考力を教えるために何をやってるんですか」と聞いたら、ある学科の先生が「みんなクラス・ディスカッション(授業内の議論)だ」と言うのです。クラス・ディスカッションで批判的な思考の能力を養成してる。表面的に聞くと、そうなのかなと思いませんか。みんなで議論し、他の学生や先生の意見聞いたりのいいですね。「それでは、先生は批判的な思考をどのように測定したり、把握してるんですか」と聞くと、あまり明確な答えが返ってきません。ディスカッションなのだから、教室内の議論を観察していれば分かるという次元の話ではありませんよ。もし私なら、例えばルーブリックを使って学生の発言を測定するとして、意見を述べる際に、その補足データを参照するだとか、反論も考慮して理論構成するとか、測定項目をいくつか考えます。

もう1つの例としてインフォメーション・リテラシー、情報リテラシーっていうのもよくアメ

リカでは一般教養の技能の1つとして謳われています。同じように、教員に「先生の科目では情報リテラシーをどのように教えてますか」と聞きます。大体、二つの答えが返ってくるんですね。

「俺はエッセーだ」とか、「小論文で教えてる」。もちろん、それはそれでいいんですけど。他によくあるのが、「うちの授業に図書館員を呼んできて、データベースの検索を学生に紹介してもらってる」から、それで情報リテラシーを教えてるという理解がされています。カリキュラム・マッピングの目的は、各教員の教授法の良し悪しを判断することではないので、まず始めは全て肯定します。「A先生のやり方は興味深いですね、どのような題材を学生に提示していますか」や、「小論文はどのような内容を学生に課していますか」、または「どのような教科書を使って情報リテラシーを教えていますか」等々、何でも良いのですが、教員が自分達の実践を気後れせずに話せる雰囲気を築いていきます。だいたい教員からの実践例が吸い上げられた頃に、「もう少し他の方法で情報リテラシーの能力を教えたり、把握する方法がないですかね」というふうに話を展開します。

さて、細かい具体例はこの辺にしておきましょう。カリキュラム・マップの目的、すなわちどんな役割を持った道具かを考えてみます。まず、カリキュラム・マップは本当に導入の導入というか、アセスメントの入り口の部分になります。このマップさえあれば、ある技能はここの辺の授業で教わってるから、ここの授業を中心に何かの共通テストを実施してみましようとか、共通のルーブリックで学生の能力を測定してみましようという次の次元の話ができる。その道先案内板みたいなものになります。なので、いくらカリキュラム・マップを作成する際に教員間の議論が盛り上がり、充実したものであったとしても、カリキュラム・マップの作成自体が学生の学習成果を測定しているわけではありません。ただ、その教員の関心を持続させる為にも、次の段階の測定の活動が求められます。

【学内資料を公開】皆さんがカリキュラム・マップを実施する際に難しいと思うのが、異なる学科間でどこまで統一性を持たせて、どこまで自由度を持たせるかだろうと想像します。私なりの進め方としては、かなり自由度を持たせてやっています。というのも、最初から縛りをきつくと、A学科はこれやってないよとか、B学科はインターンシップをクリニカルと呼んでいるなど、入力する項目の数がすごい増えてしまうんですね。そういった枝葉の対応にいちいち時間と労力を取られるのも嫌なので、まずは学科の先生方で共通する用語だとか、教育方法をまずキーワードとして設定して、それで作業を始めてもらうようお願いしています。画面を見ていただくと、一般教育の技能が左側に並んでいます。口頭伝達、数量的思考、自然科学的思考だとか、社会科学的思想、そういうのがずらりと左側の列に列挙してあり、右側の上の行の所は政治学の科目です。POSはPolitical Scienceの略で、数字はおそらくご存知かと思われますが、米国のコース番です。さて、どの科目で何をやってるのかというのを示す項目としては、EだったらExamで試験、HだったらHomeworkの宿題、DだったらDiscussionの議論というふうに、誰がどこで何をやってるのかを書いてもらっています。参加者の皆さんには大体の感じさえつかめていただければいいかと思うので、カリキュラム・マップの事例は、この辺でよろしいですか。

佐藤：よくカリキュラム・マップのイメージって、講義の内容が例えばクリティカルシンキング、要はレクチャーそのものがみたいな議論があるじゃないですか。対して、本田さんの場合は、それこそアセスメントの成果やホームワークの結果だとかを見て、クリティカルシン

キングは OK というのを判断するために、分けるということなんですかね。だから、教える内容とこっちのスキルを結び付けるやり方もあるだろうとは思いますが、これはどちらかというと、測定する方法に合わせるというところに違いがあると思うのですが、この辺議論はあったんですか。

本田：佐藤さんの言うように 2 つに分けるという発想は何となく理解できますが、私なりにはあまり気にしていませんでした。多分それは暗黙の了解としてセットになっているんです。何らかの技能を教えているからには、測定するっていうのが前提としてあります。ただ、今いただいたご指摘についてもう少し議論を進めてみると、技能の列の真ん中くらいにファイン・アーツという、芸術というのがありますね。ここは、大体空いてるじゃないですか。ここで紹介しているのは政治学の科目なので芸術に関する内容は教えていないのが普通です。ところが、看護学科の学科長と話したときに、「実は私の看護の科目では芸術に関する技能も教えてるの」といわれました。「え？ 先生、どういうことなんですか」と聞き返したら、「パワーポイントのスライドには芸術性が求められるでしょう」というのです。「でも先生、パワーポイントが美しいか、美しくないかっていうのは、もちろん成果だとか、測定の判断基準にはなっても、授業の中でパワーポイントのきれいな作り方の教材を与え、時間を取って教えています？」と聞いたら、「教えてない」って言うんですね。「じゃあ、それは入れない方がいいのではないですか」というふうにさばいたんですよ。回答になってますか。

佐藤：分かります。

本田：少し注意点を加えておきましょう。私の勤務校のカリキュラム・マップの様式を導入するのは正直のところ難しいと思います。段階を踏んで、最初に申し上げたようにまずは初級、中級、上級で表を埋めながら対話を始めて、多分議論をしてる間に、「先生、初級ってどういう意味ですか」、「中級ってどのような内容が求められますかね」というふうに教員の関心が深化していくと思います。こうして時が熟したときに、多分私の大学で実施している様式の次元に移れるのだろーと思います。「中級で僕はこういう教科書使って、こういうミニテストを導入してるんだよ」だとか、「私の科目ではエッセーでも 5 ページ書かせて、きちんと参考文献は 10 点入れるようにしてるんだよ」だとか、そういうふうな具体的な話に発展させていくのが理想です。ただ、それを最初からやるのは、ちょっと難しいかなという気がします。JABEE のようなプログラムの次元での外郭団体がアクレディテーションを強硬に進めていけば実現可能な気もしますけど。特に、文社系でやるのはちょっと難しいかなと思います。

関：ご質問なんですけど、恐らくこれ、何かのある特定の分野の一覧みたいになってるイメージですね。そうしたときに、そのプログラム全体で、仮にすっぴんって言ったら変な言い方になるのですが、これができてない時点で取りあえず何か埋めるわけですね。そのときに何かの能力、トータルで見たときに、ここは欠けてるなとか、そういうことって起り得ると思うんですけど、もしもそういうのが起こったときにも何らかの調整はして、作ったときには行うわけだと？

本田：まさに関さんの仰る通りで、理想は教員が科目の一覧をみて、どの技能が手薄になっているか、学科の教育目標との整合性は十分に担保されているかという議論に持ってけたら、

カリキュラム・マップの活動の落としどころというか、成果ですよ。その議論が無いままだと、カリキュラム・マップは単なるやっつけ仕事になってしまうので。もう1つ暴露話になりますが、学科長によって、この仕事の進め方が違います。ある学科長は、私が作成したエクセルの様式を全教員に配布して、「じゃあ、頼むから適当に埋めといて」で片付けます。こういった学科長は忙しいから効率的に仕事を進めたいとか、学習成果の活動に関心が低いとか、学科内の教員の人間関係が悪いので教育の話を持ち出すと哲学論争になるだとか、言い訳は様々です。他方、看護学科や歯学学科の先生はプログラムの次元のアクレディテーション（適格認定）が凄くがっちりしてるので、教授会できちんとカリキュラム・マップの作成に取り組みます。皆で面と向かって、1時間くらい割けばそこそこできるかな。作業に掛かる時間は学科の規模や人間関係にもよります。例えば、時間が掛かった学科では、月1の会議でおそらく30分か1時間掛けて、半年を要した所もあります。

最初にも申し上げましたが、やっぱり道具って使い方なんです。コンプライアンス（法令遵守）で、単に穴埋めしとけという程度では実効性は見込めませんが、私は最初はその程度でもいいかなと思ってます。取りあえずマップさえ出来てしまえば、将来的には学科の教員に、批判的思考を教えている科目はこれとこれなので、一般教養のアセスメントの周期が来たときに何か共通テストを導入するなり、共通ルーブリックで測定してみませんかという話に持ってくるので。この辺は学科の状況によって働き方を変えます。他は何かありますか。

小湊：いいですか。ちょっと見方の問題なんだけど、左の欄の所にいろんな技能も含めて、身に付けるべき項目が書いてあるでしょう。これって、さっき上の横のほうを見てたときに、確か政治学の授業とかっておっしゃってたけれども、例えば日本だったら、学部とか、学科によってそれぞれ教育目標が定められていて、それはいかにも抽象的なので、もう少し具体的な知識のレベルに落としたらどうなのかっていうのを表現し直してマップを作ったりします。そういう観点から見たときに、これは政治学の授業がずっと並んでるけど、あれ多分科目ナンバーですよ。数字はね。

本田：そうです。

小湊：そうすると、1年次から4年次までの科目をずらっと並べるときに、どの教育をされてるかっていうふうに見るのは分かるんだけど、この一番左側の技能の部分ですよ。これって、政治学全体に共通して求められてる技能なんですか。その辺がよく分からない。

本田：すみません。説明のときに漏れましたが、これは一般教育の技能一覧なんです。いわゆる、ジェネラル・エデュケーションのスキルなので、それは専門分野、専攻にかかわらず、学生が基本的に身に付けるべき能力と規定されています。なので、左側の技能はこのセットでどの学科にも対応させています。

実は、もう一つ学科の教育目標に沿ったマップがあります。これは一般教養で求められる技能ではなく、政治学科の学習目標で明示されている技能に沿って、同じように学科内の科目が学科の学習目標と整合性が取れているかの点検をします。

あとは、どうしようかな。カリキュラム・マップについてあまり深みに入っていくと時間が……。もし行き過ぎていたら、寫田さんと浅野さんに止めていただいても結構です。実は、この一般教養の学習成果の11分野の下に、各分野ごとに指標が5つぐらいずつ揃っ

ています。というのも、批判的思考を教えていない科目なんて大学ではありえないというか、誰でも教えています。その割には、あまりに抽象概念すぎて、学科間だけでなく、学科内ですら、批判的思考の能力を定義するのに合意できないわけです。例えば、批判的思考の中でもデータを評価できる、データを解釈できる、異なる A と B の観点を検討できるといった指標が批判的思考の分野に付随しています。必要であれば、その指標を参照に各教員がどのような授業を実施しているのか話を進めていきます。

小湊：そうすると、そういう指標ってどうやって作ってます？ 学内では。

本田：適当って言っちゃるとマズイですよ。まずは、ざっくり作るようにしています。

小湊：多分そこは結構苦しむんですよ。

本田：一般的には、学習成果は教員の業務範疇ですから、下からの吸い上げが基本的な進め方なので学科の先生に投げちゃいますね。例えば批判的思考であれば、哲学科の先生とかは結構そういう内容を押さえているので、まずは下書きしてもらい、他の先生からの意見を聞きながら調整していきます。幸い、アメリカでは先行事例が多いので、他の大学や適格認定、学術協会などが整理している批判的思考の定義や指標を参考に、自分の大学に合った指標を設定するのが普通です。

小湊：難しいのは、例えば細かく分かりやすい形で技能を分解して要素に分けていくとしたときに、その要素の幾つかが例えば計測不能なものが上がってくるっていうことがよくあるんですよ。そういったときに、例えばアセスメントの立場から何か助言したり、一緒に何か共同作業をやったりすることってあるんでしょうか。それはもうそのまま置いとくものなんですか。でも、計測不能なものを置いてしまうと、結局そのアセスメントのときにそれを計れないって話になるから、その目標のところに掲げてても実質意味がない話になってきますよね。教育の中身としてはいいのかもしれないけれども、アセスメントの立場からすると意味のない項目になりますよね。そうでもないんですか。

本田：多分、小湊先生のおっしゃってるのが、機構の田中先生みたいなエバリュアビリティって意味でのご質問であれば、学習成果だとか、学習目標って抽象度が高いものから、どんどん小項目に具体化されていきますよね。例えば国際人の養成ですが、国際人ってどういう意味ですか？ すごい典型的ですけど、英語話せる人とか、海外情勢に詳しい人なのか、英語で仕事が遂行可能な人なのか、そういうふうに異なる側面を小項目に落としていくので、測定不可能なものは細分化していくよう誘導します。もし、抽象的な定義や指標が示されたら、「もう少し具体的にご説明いただけませんか」っていうふうな議論で落とし込んでいきます。

ただし、問い詰めすぎたり、完璧主義に陥らないように気をつけますね。先生があまりに問い詰められて、疲れたりだとか、逆に「本田、うるさいな、もうしつこいよ」と言われる限界点があると思うんですよ。そこにいく前に、「ありがとうございます。じゃあ、今の妥協点としてこういう定義で載せておきましょう」と一歩引き下がります。そして、何年後かの学部の自己点検のときにどう測定するか、またあらためて議論しましょう」と言って締め括ります。私は争いをできるだけ避けるように心がけているので、そういうふうにはやってます。まとめれば、妥協点を見つけるっていうことですかね。お答えになってます？ これ。

小湊：いや。それは多分アメリカの教育の体系と日本の学部ごとにあったりとかいうのはちょっと一概に比較しにくいので、何とも言えないけど、同じことはやっぱりうちの大学でも起こるわけですね。例えば国際人を養成するっていったときに、じゃあ、何をもって国際人と判断しますかって話は当然します。そこで、いろいろ分解して出てくる中で、およそこれではちょっと何をもってそれを根拠にして出せばいいのかが分かんない項目って、やっぱり出てくるんですよ。

本田：ありますね。想像はできます。

小湊：そこはもう置いておかざるを得ないと。ただ、そのときにさらにアセスメントの立場から議論を深めるために何かファシリテートするのかなっていうことで質問しているんですけどね。

本田：先ほども申し上げたように、多分カリキュラム・マップ作成の時点ではあまり厳密にやり過ぎないようにしています。もちろん、アセスメントの教科書や適格認定の教科書では学習成果の定義から作業を始めるように推奨しています。ただ、データもなしに観念論だけで議論しても堂々巡りなんですよ。

カリキュラム・マップって、あくまでも誰が何やってるかという確認の為の道具です。例えば、実際にインフォメーション・リテラシーっていうのが下から2番目にありますけど、これを去年測定したんですね。業者のテストを入れて。そのときに、サンプル取るのはPOSっていうポリティカルサイエンスの234と405からサンプル取りましようっていう理由付けとか、糸口にはありますけど。もちろん業者のテスト入れたから、測定できてるとか、そこで教えてる先生が納得いくようなことになるかというのは別の次元の話です。

小湊：そうですね。

本田：納得しない先生は、そんな Multiple Choice っていう何て言うんですしたっけ。多肢選択？

浅野：複数選択または、多肢選択ですかね。

本田：「それで計れる能力なんて、大学教育じゃない」みたいなことを言う人もいますから。それに対し、私は「でも、先生、観念論で言っても共通理解が進まないの、まずやってみましょう」と言って、だましまし進める。そして、データが取れたときに、あらためて議論しますね。そのデータ取る前に議論すると、それだけで何年も費やしちゃうので、まずはデータを取らせてください。不完全でもいいから取ってから議論する。

浅野：確認も含めて質問させてください。いま、お話しいただいているカリキュラム・マップを作成するときに、General Education (教養教育) っていう話がありましたけど、カリキュラム全体で見たときに、教育目標を変えることも考えていく必要がある場合があると思うんですけど、その場合、どう対応されるんですか。

本田：もし教員の方から一般教養の分野と各科目の目標や教育活動に齟齬があるという節の発言が出てきたらしめたもの。「先生、それは素晴らしいご指摘です」と伝え、「では、学習目標の見直しも学習成果を測定する上で大変意義のある活動なので、学習目標の再定義をしましょう」と提案します。だから、私の仕事って何でもかんでもデータ集めて測定するだけじゃなくて、測定手法自体を変えたりだとか、目標の再定義っていうのをよくやります。

浅野：日本の大学では、多分、ここでつまづいている場合が多いと思います。したがって、ここでつまづいてしまうと、後の測定がうまくいかないということになります。そのため、私は学内でここに時間をかけてやりましょうというふうに働きかけています。ただ、いまのお話をお伺いしていると、結局、最初はあまりがしがしやり過ぎると時間だけかかるので、ひとまず走ってみて、測定をして議論をし直していく、そういうことですね。

ご質問者：質問よろしいでしょうか。そちらで今、能力要素ごとに下位項目も作られているとおっしゃったんですけど、それに対するルーブリックっていうのは、それぞれ対応して作られているのでしょうか。

本田：あります。それは、例えば表の一番上にあるWritten Communication（文章作成）だと、アメリカでは文章作成能力のルーブリックはおそらく何十個もの例があり、簡単にダウンロードできます。それを2つ、3つ見て、自分の大学や学科の教えてる内容に合ったようなルーブリックを作ることは、アセスメントを比較的しっかりやっている大学ならやっていると思います。

補足：ただし、米国の私立大学や研究大学はそこまで出来ているか定かではありません。話は変わりますが、日本の状況を傍目で見ると、文科省がルーブリックといったから、測定方法として採用するだとか、ある大学ではルーブリックがあたかも魔法の杖のように喧伝されていますが、測定方法の手順をきちんと行っているのか不安です。私は学科の次元のルーブリックによる学習成果の測定では最低でも1つの小論文を2名の教員に査読してもらい、信頼性の検定を掛けています（inter rater reliabilityとしてintra class correlationを使っています）。

一般教養で身に付けるべき技能は、その性質や教員の関心によってルーブリックを使ったり、共通試験を使ったり、そうでなければ学内で自作の小テストを作ったりと、測定手法はその時々で変えていきます。例えば、今回の情報リテラシーは業者のテストでやってみたけど、学生が情報を評価する技能は複数選択では測定しきれない。であれば、次回の情報リテラシーの診断では小論文をルーブリックで測定してみましようということもあり得ます。

いろいろあると言うとあまり答えになっていないと思うので、もう一步踏み込んでお話しすると、私は診断業務の着地点はお膳立てはするけど、判断はしないんですよ。最終的には教員の業務ですから。「先生、今回の診断ではルーブリックでなければ絶対駄目です」なんて言いません。よくあるのは、教授会でだいたいルーブリックでいいだろうと集団の合意が成立しつつあるときに、私が専門家として承認すれば後押しにはなりますけど。教授会内で意見が分かれているときに、私独自の判断である手法に決めてしまうことはありません。

人間ってやっぱり経験しないと分からないですから、まずは教員がやりたいようにお膳立てして、測定の活動を実施してみることに。その後に、「あのとき本田がルーブリックがいいよと勧めていたけど、やっぱりそうだったのかもしれない」と1人でも2人でも分かってくればいい。次に実施するときには別な手法でやってみようかという話になると思う

ので、取りあえずは経験してもらおう。

小湊：ちょっと踏み込んだ質問なんですけど、次やる機会、何か決まったサイクルがあるんですか。それとも、割とアドホックにやるものなんですかね。つまり、うちの大学では、それが無いので、決まった契機みたいなのが無いので、1回やるとやりっ放しで終わってしまう。だから、改善に結び付かないってパターンが出てきますよね。

本田：それに関しては、学習成果の診断（アセスメント）の周期があります。学内の資料をご覧くださいませ。これが一般教養教育の診断で5年周期です。これは、3年周期にしている大学もありますが、私の以前の勤務校がそれで、データ集めに追われて教育改善を議論する時間が取れませんでした。なので、現在の勤務校では5年周期に延ばしたのです。この周期の長さは、他大学を参照にしながら、機関適格認定を意識して学内の合意形成で決定したものです。

これとは別に、各学科の自己点検、プログラム・レビューがあります。こちらは州立大学機構の取り決めで5年に一度は自己点検をするよう定められています。

大体、アクレディテーションと呼ばれる適格認定で一番先に見られるのは、きちんとサイクルできてるか、それに沿って診断が進んでいるかだと思います。少なくとも、米国北東部のニューイングランド地域の機関適格認定は、そこが基本線になっていました。一般教養の診断の周期に関してですが、少し自慢話をしてもいいですか。

小湊：どうぞ。

本田：実は今お見せしている一般教養の診断の周期は機関別適格認定で結構いい点をもらいました。というのは、多くの大学が診断の活動を単年度でしか考えてないんですね。1年でデータ集めて改善までつなげるのは時間的にかなり難しいでしょう。ですから、私は1つの診断活動を3年掛かりで設定しました。1年目が試行段階で小規模で測定してみる、2年目が本番で、試行段階で得た情報から測定手法を改善し、大規模に実施する。3年目は前年のデータ分析をもとに授業法やカリキュラムを変える。米国ではこれをClosing the loop、いわゆる結果の活用と呼んでいます。実際は結果の活用は多くの大学で出来ていないという報告や研究が山ほどあります。そして、私の大学でも周期を設定して走り始めてまだ3年目なので、結果の活用まではあまり進んでいません。

この周期を設定する際に苦労したのは、一般教育のアセスメントのサイクルと、学科の自己点検のサイクルの調整です。この表に、SSとERってありますね。SSというのはセルフスタディーで自己点検、日本も一緒ですよ。次年度にエクスターナル・レビューといって外部者による実地審査があります。これは国立大学でもやっていますよね。各学科のプログラム・レビューに2年掛かるので、それに重ならないよう一般教育のアセスメントのサイクルを組んでいきました。

小湊：これとは別途アクレディテーションのサイクルを当てはめるのもあるんですね。

本田：あります。基本的には実地審査は10年に一度ですが、その間に5年目で中間報告書を提出する必要があります。ただ、実地審査で勧告を受けると、内容によっては2、3年のうちにもう一度、課題を絞った小規模の実地調査が入ります。

小湊：そこが多分一番肝心ですね。

本田：確かに、機関別の適格認定とのすり合わせは重要ですね。学内の学部学科レベルの自己点検と外部審査だったら、来年度はこの学科とあの学科の順番を入れ替えることもあります。例えば、学科長が替わったとか、転職して自己点検が頓挫したりが理由です。でも、機関別のアクレディテーションに対応できてないぞという次元の問題は学部長や学術担当筆頭副学長（Provost）と調整して進めています。よろしいですかね。皆さん、眠くなってないですか。大丈夫ですか。もう2時になったので、あと1時間しかないのですが、まさかカリキュラムマップでここまで引っ張るとは思わなかった。

小湊：ごめんなさい、ちょっと突っ込んでしまった。

本田：いえいえ、大丈夫です。おそらく、日本ではまだこの段階が大事なのでしょうから。ただ、私の講演者側からの勝手な意向だと、もっときらびやかなデータ分析ってあるんですけど。そういう話よりは今のところ学習成果の診断の基礎固めのところになってますけど、いったんここで休憩入れますか？

浅野：このまま続けて、いいんじゃないですか。これまで、いろいろ質問が集中しているのは、アセスメントをどのようにやるのかということのように思えます。本田さんがおっしゃるように、ルーブリックとか、ツールはいろんなものがあるのは分かってるんですが、やっぱり教育目標、カリキュラム・マップの作成をしっかりとやらないと、結局、後でつまづくって話になるんだと思います。

小湊：やっぱり制度設計って、手順をどう組んでいくかって結構大事だと思っていて、国からはこういうの作れって当然言われます。作ることは簡単なんです。ある意味。作文すればいいから。

本田：そうなんですよね。

小湊：だけど、それじゃ何のためにやってるのかってところでみんな疑問に思うわけですね。これだけ時間かけてやるんだってやっぱり思うから、どうやって改善をうまく回せるようなサイクルだとか、仕掛けをつくっていくのかなってというのはやっぱり常に気になってるところなんですよね。

本田：先ほども申し上げましたが、米国でも同じようにカリキュラム・マップをやっつけ仕事で片付けてしまう学科もあります。ただ、それだと小湊さんが仰るように、教員が何の為にやっているのか分からないまま仕事が増えて怒るわけです。なので、そのうち2、3分の短いビデオでも作って、教員にカリキュラム・マップの意義を周知したいなと考えています。仕掛けについては、やはり診断の周期と測定の設計が鍵を握っています。

GPAの活用状況(1)

- 奨学金、卒業基準、表彰などで活用
- 適格認定の学習成果の基準では活用不可
 - 教員間による成績付けのバラつき

12

本田： さて、GPAに話を移しましょう。学内の例は後でお見せするので、まずスライドの部分をさっとお話ししますが、成績分布でデータを見ます。

GPAの活用状況(2)

- 難関科目 (Barrier Course)
 - 学内の文脈
 - データ分析例

参照例 <https://www.fgc.edu/wp-content/uploads/2011/12/statbriefmar2008b.pdf>

13

スライドに難関科目とありますね。GPAを使って、より正確に言えば成績分布を通じて難関科目を見つけます。例えば、社会学のソーシャルサイエンスのSOCの305と適当な番号を当てましょう。科目名は社会科学の研究手法。どうやら、そこで多くの学生がつまずいてるらしい。どこの学科にも難関科目があります。米語ではバリアー・コース

と呼ばれるのが一般的です。

さて、もっと泥くさい話に入っていきますよ。バリアー・コースって、成績の甘辛でいったら、辛い点数付ける先生が教えてたりします。結局、「原因はそれかよ」と誤解されがちな話なんですけど、その先生を説得して成績インフレを起こせばいいのかというわけでもありません。例えば、私の知る限りでよくあるのは、難関科目には100番台と300番台が多いんですね。なんで100番台が多いかというと、1年生の経済学原論ってあるとしますよね。突然高校生から大学生になってマクロの経済のいろんな理論の話されても分からないと思います。特に、こうした科目は必修で大体は大講義ですから、そこで科目を落としてしまう場合が多いわけです。だから、100番台で心理学概論、経済学概論、何とか概論みたいのは危ないです。

もう1つは先ほど申し上げたように、300番台です。これは専攻科目の入り口で、本格的に専攻の技能を身に付けなければならないところ。工学や生物学などのハードサイエンスは一般的に成績が厳しいのと、他では数学や、社会学とか、心理学は技巧的な部分が割合大きくなるので難関科目に特定される場合があります。

データ分析ではどの学科のどの科目がバリアー・コースか把握できています。また、全学平均でいうと、各科目でAからC-までを米国ではコース・サクセス・レートと呼んでいます。それが、学期にもよりますが、75パーセントくらいです。日本では退学率という見方が一般的なので、それに置き換えれば、25パーセントがFかドロップアウトしちゃいます。このコース・サクセス・レートが低いところだと、45パーセントぐらまで落ちます。

では、この難関科目に対して、どのような対策を取ったのでしょうか。教学担当の筆頭副学長が何をやったかという、年に2回ほど全学学科長会議というのを設けています。そこで、各学科長に難関科目の見直しを迫りました。具体的には、コース・サクセス・レートが50パーセント台の科目を洗い出し、その中から改善の優先順位が高い科目を1つに絞り、原因がどこにあるのかを考えます。簡単に言えば、教育の仕方が悪いのか、成績の付け方が厳しすぎるのか、先修要件の整合性が整っていないのか、などがあります。その辺をきちっと押さえてくれというふうに学科長に投げて、改善案を持ってきてもらいました。それが1年ほど前の話で、また2、3年後に結果が変わったのかを検証することになるでしょう。

GPAの活用状況(3)

- 学習コーチ
 - 学内の文脈
 - 成績分布の分析
 - 集団面接による質的情報の分析

14

もう一つ、別のGPAの活用事例についてご紹介します。簡単に言えば、難関科目に学習コーチを導入した効果の診断です。

勤務校は社会人学生が多く、全学平均で33歳くらいです。こうした学生は常勤の仕事に加え、子守りや年老いた両親の世話などに追われ、学業に付いて行くのが大変なことがあります。

そこで、学習コーチを入れて、学生が毎週きちんと宿題を提出してるかを確認し、未提出の学生には連絡を入れて、何とか授業について行けるようにお世話します。状況としては、ここ数年でオンライン授業の数がだいぶ増えました。オンライン授業の難しさは、学生があまりパソコンを使えないのに履修登録してしまい、いざ授業が始まってから課題の提出の仕方が分からないことがあります。

Blackboardって日本にもあるのかな。

小湊：ありますよ。

本田：Blackboardというオンライン授業のポータルで、課題の提出の仕方が分かり難いときがあります。私自身、Blackboardを使ったオンライン授業の教え方の科目を履修したのですが、正直、自分の回答が提出できたのかなって分からないときがあるんです。こうした問題に直面した学生を助けるために、学習コーチを入れました。

その学内データをお見せする前に、最後の適格認定の所だけお話ししておきましょう。機関別適格認定では、GPAは学習成果の指標としては使えません。というのは、これはベミジ州立大学の藤原さんも発表なさっていましたが、直接指標と間接指標という考えで、成績は間接指標なんですね。なぜ間接なのかというと、もう皆さんもお分りのように、先生の成績に甘辛があります。例えば、私が統計学を取ったとして、あまり統計の試験で点数がよくなくても、授業はいつも出て、宿題やって頑張ったからBをくれる先生がいるかもしれない。そういう自分が習得した能力とは別に、授業の出席や議論への参加で点数を稼ぎ、本当は学習成果としてはあまり能力身に付いてない場合があります。そういうのを考えると、適格認定でGPAっていうのは学習成果の証明にはならないわけです。

[学内の報告書を公開] 話をもとに戻します。さて、この学習コーチの導入の成果を測定するにはどうすればいいのか。まずは、今回私自身がデータ分析で学んだこととお話します。最近の米国高等教育界ではスチューデント・サクセス（学生の成功）というのが流行り言葉です。私はもともとIR部署で仕事をしていたので、ここでのサクセスの定義は、AからCマイナスなんですよ。だから、その定義でデータを分析したら、学習コーチの効果が顕著に現れませんでした。紺色の表が去年のデータと過去3年のデータです。平均サクセス・レートで言えば、62パーセントとか、80、74、75。これが学習コーチ導入前です。それと比べて、上の緑色の表が学習コーチを導入した今年の夏のデータで、66パーセント、71、86。ぱっと見、あまり変わらないんですよ。ヤバイ、学習コーチの効果が無いぞ。上にどうプレゼンしたらいいのかなと考えました。そこで、まずはエンロールメント・マネジメントという、学生の募集から在籍管理を扱う部署に行って、そこのディレクターやスタッフと話をしてヒントをもらいました。「本田、授業をドロップ・アウトしちゃう学生というのは、AからCじゃないから、Dの成績も見てみてよ」と頼まれました。WというのはWithdrawで、科目を落とした学生です。そのWithdrawのWとDを見ると、昨年紺色の表と今年緑色の表を比べると、Dの割合が上がってるんですよ。具体的には、2014年の夏にはDもらった学生が3.5パーセントだったのが、今年は8パーセントに増えている。

IRの観点からすると、「それってスチューデント・サクセスの定義外だから」と見落としていたわけです。でも、エンロールメント・マネジメントの観点で見直すと、「Dっていうのは大切なんだ。学生っていうのは、最低でもDさえ取れていれば、連邦政府の奨学金を受給する権利を維持できるから、退学率をこれで抑えられるんだ」と教わりました。これは分野が違うとデータの見方が違うので、目からうろこという感じでした。すごい単純な話ですけど、こうやって他の部署の人と議論を重ねてデータの分析や作表を仕上げていくという事例でした。

浅野：この表でいくと、Dが増えていますが、逆にWは減っているってことですね。

本田：そうなんです。

浅野：放棄してる学生が減ったっていうことですね。

本田：放棄してる学生が減って、何とかみんなDで食いつないでいる。奨学金受給権を維持してるってことは、多分リテンション（在学維持）につながるんですね。来年もその学生は恐らく来てくれるだろうっていうから、顧客としてはきちんとつないでおける学生。GPAの使い方は、このように成績の分布を見てどうするかって話なんですけど、これまでの話で何か質問やコメントなどありますか。

鳥田：表の見方を。

本田：もうちょっと丁寧にやったほうがいいですかね。

鳥田：結局、各成績のAからFまで、Fが不合格で、Dまでが合格？

本田：そうですね。（画面を変えて）これだともうちょっと見やすくなりました？ 浅野さんがおっしゃったように、Wが減ってDが増えたっていう、それだけなんです。落としどころは。ただ、その意味は、エンrollment・マネジメントのディレクターからすれば、意味が大きかったわけです。

鳥田：DはGPAとしてはグレードポイントは何点なんですか。

小湊：1でしょう。

鳥田：1ですね。

本田：だから、複数の科目でDを取り続けるのは駄目なんですけど、Dを1科目、2科目取るのは、他の科目で頑張れば何とかGPAの点数は戻せるんですよ。

鳥田：要するに2.0が進級要件みたくなくて、だから、Dを取ってるってというのは要するに取れば取るほど（進級的には）きつくなるけども、要するに奨学金とかを維持する上では一応授業は取ってるよねっていうことになるっていう。

本田：まさにそうです。

鳥田：そういうのが、Dっていう意味で。

浅野：日本の考え方に置き換えると、Aのところが多分、多くの大学で設けている「秀」で、その後が「優」「良」「可」「不可」だと思うんですよ。表のFは、不可。だから、Dが可ってことですね。

本田：何となく僕、これ、いいネタだなんて思って持ってきたんですけど、皆さんの顔がちょっと、「ん？」みたいに・・・。

浅野：いいえ。多分鳥田さんが言うように表の見方です。多分そこが分かれば、皆さんにもご興味のあることだと思います。

本田：そうおっしゃっていただくと救われます。

鳥田：1行、行は何？

本田：行は科目です。ビジネスの101、CISはComputer Information Scienceの101と続きます。先ほども申し上げましたが、一般論としては100番代の科目は新入生が大学の授業についていけなくてサクセス・レートが低い場合があります。ただ、今回のデータで注意していただきたいのが、今年の夏はあくまでも学習コーチの導入段階だったので、この試み

に賛同する教員を募って試行しました。なので、本当に学習コーチが必要な難関科目には導入していません。学習コーチの導入に理解のある教員の科目はもともとサクセス・レートの高い授業だったので、正直のところ伸びしろが少ないんですよ。来年辺りから、もう少しサクセス・レートが低い科目を対象に導入していく予定です。

なぜ、この事例が面白いのかというと、2つあると思うんです。1つは、学生が奨学金の受給権を維持できるっていうこと。やっぱり、大学のビジネスと直結してきますから。もう1つは、IRとEM（エンrollment・マネジメント）でデータの定義だとか、見方が違った。それをきちんと解釈したり、コミュニケーションを取って、この表は簡単に見えますけど、ここにいくまでに2、3回はやりとりしていますから。もし、私がDの分析をしなかった過去のバージョンで財務担当副学長にデータを出してたら、このプログラムがつぶれるとまでは言わないですけど、来年の予算は減らされていたかもしれないですよ。それを、EMのディレクターとお話しして、「ここの項目のDを入れてよ」と表を作り変えてみると、「何だ、ふたを開けてみたら、成功の兆しがあるじゃないか」と学内の受け止め方が変わったわけです。

最後にもう1点だけ加えます。学習成果っていうと、みんな上を見るんですね。「うちの大学から優秀な学生出すぞ」って思うのが普通なのかもしれません。ところが、今回の事例は学習コーチに拾ってもらえる学生、ある意味下をすくい上げる意味での学習成果というか、学習支援っていう形って、あまり他大学で事例がないでしょう。そういった意味では、面白いと思うんですよ。いつも上を目指して、もっとできないと駄目だっていうんじゃないくて、できない子をどういうふうに支援できるかっていうのが、表現できる事例なのかなと思います。

浅野：現状、日本の大学ではこれはできていないように思えます。日本では今、リベラルアーツ教育とか、初年次教育の重要性がいわれているわけですね。一方、この表が意味しているところは、本来、学生が次のステップに進むとき、必修科目があって、それをつまずくと次にいけないわけですから、それを一覧にして、どこが学生さんのステップアップを妨げるかっていうのをしっかりチェックをした上でやっているっていうことですよ。だから、そこの意味での学習成果っていうことになるわけですよ。

本田：そう言っただけだと嬉しいです。では、この事例についてはこの辺でよろしいでしょうか。

藤原：最後、1点だけ質問を。先ほど適格認定のお話で、機関別の適格認定ではGPAはあまり学習成果の証明にならないというお話があったと思うんです。それで、今、日本では学生の履修成績のデータ、GPAなんかを学籍番号とかでひも付けながら、あとは学生調査ですね。アンケートの結果とひも付けて、実際に今いる大学の中の学生がどういう層なのかというようなことを一定分類しながら学習支援をこういう層の学生にはこういう支援をしていこうみたいな、そういう取り組みが幾つかの大学で試みられているんですね。この辺りは何かアメリカでも似たような取り組みがあるのか、あんまりア krediteーションでは評価されないんで、そもそもGPAがですね。なので、あまり取り組まれてないのか、ちょっとその辺りを教えていただけるとありがたいです。

本田：ご質問への回答の3分の1くらいは、次の学生調査の部分で出てくると思います。結局、学生調査は自己申告で間接指標ですから信頼性と妥当性がないって言われてしまうんですね。自己申告でいくら自分の批判的思考の能力が高いと思っても、それがGPAと正の相関にあるとは限りませんから。ただし、学習成果のデータ分析でGPAがしばしば使われるのは確かです。その際は、従属変数としてではなく、独立変数としてですが。IRの分析でよくあるのは、GPAを使って進級や卒業を予測するものです。だいたい、GPAが高い学生は進級と卒業する割合も高いという当たり前の結果が出ますが。あとは、うーん…。無理に分析してみるなら、学生調査で学生の学習経験の頻度や良し悪しを独立変数とし、GPAを従属変数にもできなくはないですが…。こうした分析はあまり聞いたことが無いし、自分もあまりやってみようとは思いません。話を藤原さんのご質問に戻れば、今ご提示いただいた日本での取り組みっていうのは、なかなか面白いと思います。GPAと学習経験をひも付けして、学生をGPAによって4つくらいにグループ分けして、学習経験や満足度などをANOVA（拡散分析）で差があるのか見てみるとか。そこまでしなくても、単に記述統計で下位の学生層に際立って好ましくない傾向が見られれば、そこを支援するという方策も打てますね。

もう1つ、最近私が関わってるのに、学生アドバイジング（履修登録などの相談業務）のツールがあるんですね。それだと、まだGPA使ってます。GPAが2以下になったら、確か大野さんの大学で早期発見のシステムとか作ってましたよね。私の大学で導入準備しているシステムだと、学生のGPAが2.0以下になった学生を赤で表示して、教員や学生履修相談の専門職員が相談に当たります

日本の講演で「GPAはアメリカでどのように活用されていますか」という類の質問を受けて答えにくいのは、アメリカで使ってる所と使えてない部分とすごい入り組んでるんですよ。学習成果では使ってないけど、学生の履修相談（Advising）やリテンション（在学維持）の分析では使っています。ごめんなさい。中途半端なお答えになってるかもしれないですけど。

小湊：補足になるかどうか分かんないけど、要は連邦が奨学金受給を学生に認定するために、要するにSatisfactory academic progressで2.0っていうのが連邦から示されているから、各大学は基本的に累積GPA、または期ごとのGPAが2.0を超えていないと、その学生が奨学金をもらえなくなる。だから、そこをきちっと把握するために、やっぱりGPAを分析して見ているっていう、そういう見方でいいんですよ。

本田：まさにそうです。

小湊：多分そこが日本とやっぱり随分違うなと思うんですね。日本はそういう仕組みないですから。

本田：でも、日本も奨学金は平均評定を取ってまいせんか？

小湊：GPAとちょっと違います。大学によって、そこの判断が違うのかな。

佐藤：学内の奨学金とか、給付金のときにはGPAを使ってるっていう。

小湊：そう。

本田：では、このGPAの話はこの辺でよろしいですか。今の時間が2時20分で、あと学生アンケートもありますが、講演を止めて、もう質疑応答に入ってしまうか。

学習実態調査

- 全米学習実態調査
(National Survey of Student Engagement)
- 質問項目の構成
http://nsse.indiana.edu/html/survey_instruments.cfm
<https://jsaap.jp/index.html>（日本語の対照版）
- 学内の文脈
- 経年分析の例
- 全学目標の設定

15

浅野：学生調査か学習実態、どちらかの類型を一つ、お教えいただきたいと思えますけど。

本田：わかりました。では、スライドは飛ばして、何をやったかだけ簡単にご説明しますね。

日本で広く知られている学生調査は、山田礼子先生のなさってるアンケートだと思います。あれはアメリカのUCLA（カリフォルニア大学ロサンゼルス校）の

全米学生調査を訳したものです。それと似たような取り組みでは、国際基督教大学がインディアナ州立大学の全米学生調査を訳して実施したようです。こちらは山田先生のプロジェクトとは違って、個別大学の単発的な取り組みだったような記憶があります。

小湊：項目としてはいろんな所で訳されていたり、部分的に使われたりはするけど、全国的な調査としてNSSEみたいなものはないです。

本田：そうですね。まずは私なりの見方ですが、アメリカの学生調査の設計として、学生の学習経験と学習成果、そして学生の満足度の3つの項目を組み合わせるのが一般的です。学習成果の点ですが、学生本人から見てどういう技能が身に付いたかという自己申告なので、間接指標となります。

【学内の発表スライドを公開】

これが、今年9月に発表したものです。プロジェクトのチームとしては学生部の部長、年配の生物学の学科長、社会福祉の学科長に参画してもらいました。このスライドは学内だと30分かけて説明したのですが、今回は時間の関係で簡単などころだけ、掻い摘んでご説明します。

このデータは過去、2007年、2010年、2013年に実施された学生調査の経年分析です。NSSEという全米学生調査の対象学生は1年生と4年生です。なぜかという、1年生の回答と4年生の回答を比較して、どれだけ学習経験や自己申告による学習成果が変わったのかなどを見るわけです。厳密に言えば、同じ学生を1年生と4年生のときにアンケート調査しているわけではないので、パネルデータではありませんが。

次に、学内の背景についてお話しします。私が分析で関わったのが2014年で、前2回の学生調査は学生部長と生物学の学科長が実施しました。その方達の功績は、結果を基にアクション・プランを作ったことです。正直、アメリカの大学の多くは、NSSEの実施本部のインディアナ州立大学が作成した学生調査の結果を自分の大学のウェブサイトに載せているだけです。まあ、学生調査の参加費として数十万円を支払っているだけあり、基本的な記述統計と他大学との競合分析ま

でやっておりますから、見栄えはいいんですけどね。

話は勤務校に戻りますが、過去の2回の学生調査、2007年と2010年の結果を元に改善の取り組みを進め、2013年のデータを分析してみたら、実際に効果が上がっていました。具体的には、米国ではキャップ・ストーンと呼ばれる科目があります。4年生が卒業前に履修すべき科目で、今までの教育課程で学んだ内容の集大成として網羅するのが一般的です。日本でいうと、卒業論文に近いでしょうかね。そのキャップ・ストーンの授業を取っているかという質問で、ご覧の通り、2013年ではだいぶポイントが上がっています。実はタネあかしなのですが、勤務校では以前からキャップ・ストーン科目を設置している学科があったのです。それが、全学科まで広げたことが1つ。もう1つは、いくら教員がキャップ・ストーン科目を教えていても、学生の方はその用語すら知らずに、4年生の最後に履修する科目としか理解していませんでした。そこで、教員達が「キャップ・ストーン科目」という用語を意図的に学生に伝えることで、学生がこうした授業を履修しているという回答率が上がったわけです。

次の事例は、学生のキャンパス・イベントへの参加率です。勤務校は学生寮が無く、社会人学生が通学、若しくはオンラインで履修する大学です。日本では学生が大学に通学するのは普通ですが、アメリカでは学寮があり、多くの学生がキャンパスで生活する大学の方が普通なのです。

通学制が学寮制かという環境の違いで、学生の回答に大きな違いが出るのが大学イベントへの参加です。当然ですが、後者に有利な回答が出るのは明らかですね。勤務校では2007年と2010年の調査では学生の大学イベントへの参加率が低かったそうです。その対策として何をやったかという、お昼休みをつくりました。12時から1時は全学でお昼休みにして、学外の専門家を招いた講演会や学内の教員の研究発表、健康や政治参加など、様々な分野の催し物を増やしたら、学生の参加率は2013年の調査で上がりました。

その他、全体的な傾向をお伝えします。まず、1年生は緑の棒グラフを見ていただければ分かるかと思いますが、学習経験が良い方向に変わってきています。対照的に、4年生は自己申告の学習成果の分野で複数の項目が減少していました。要するに、筆記や口頭のコミュニケーション・スキルや他者との協同する力があまり身に付いていないと学生が報告しています。

学生サービスの面では1年、4年生ともに満足度が高く、これらの数値は競合校と比較しても勝っています。

経年分析の手法ですが、ANOVA（分散分析）を用いました。NSSEの分析ガイドで勧められていたのですが、まずはANOVAで学生調査を実施した複数年の平均値の差に統計的に有意差があるかを確認し、Post-hoc test（事後検定）としてTukey HSDを使用しました。簡単に言うと、2007年、2010年、2013年の平均値を取って、その中でどこか1年が平均よりも統計的に有意差があるかどうかを検定します。ただ、単に有意差があるかだけでは、その差の規模が分からないので、エフェクト・サイズの検定としてOmega Squaredを使用しました。

正直、この辺の手法の話は学内の発表でも社会学の教授が批判してくる際の守りの情報であり、執行部がここまで手法を理解する必要性はあまりないと思います。もちろん、結果はしっかりと認識してもらうのは大事ですが。

グラフの見方に移ると、緑色が1年生で紺色が4年生の結果です。棒グラフ白抜きしてあるのは経年で統計的有意差が無かったことを表しています。対照的に、棒グラフが緑か紺で塗りつぶされていれば、統計的に有意差があったことを示しています。グラフにオレンジの枠が示されてい

る場合は、エフェクト・サイズ、いわば経年の差が中 (0.06) 以上に大きかったことを示しています。

統計分析にはSPSSを使いました。日本にもありますよね。SPSSでグラフを作れるのですが、それだとバーグラフの上がり下がりが物凄く大きく出るように設定されています。X軸の最高値と最小値の設定の問題なのですが。このままでは間違っただけの印象を受け手に与えてしまうので、私なりにエフェクト・サイズに見合った形でX軸を再設定しています。

私が精力を注いだのは、いかに1枚のスライドに複数のデータ項目を簡単に理解できる形でまとめるかです。このグラフの色付けの設定は自分なりに考案しました。目的は、執行部や教員にどの分野の改善を促せばいいのか、その議論の基礎資料として、学内議論の焦点を浮き彫りにする為です。学内では既に自分の先入観に捕らわれた方々も多いので、分析結果を見て気付きを得るか、それとも我田引水的な解釈をするか分かれるところですが。ただ、グラフの表示の仕方はとても分かりやすいと、学内外のIR関係者には褒めてもらいました。

IRの専門家になるには統計技能が必要だと考えていらっしゃる方々が多いかと思います。確かに、最低でも t 検定や重回帰分析、ロジスティック回帰分析はIRの基本なので身に付けておくことをお勧めします。ただ、ベミジ州立大学の藤原さんのように高度な統計分析が求められるかといえば、その機会は限られていると思います。大規模な研究大学（私なりの判断では学生数が1万5千人以上）で7、8人の大所帯のIR部署だとSenior Analystが置ける財政的な余裕があり、統計の専門家が居ると高度な分析ができて大助かりです。ただ、小さな2、3人のIR部署でそれが難しいでしょう。正直、私も藤原さんの統計のレベルまで少しでも近づけたらいいなと羨ましく思っています。

今、32分ですか。じゃあ、急いでこれをお見せしたら、いったん質疑応答に移りますので。あれ、アットマークってどこだろう。ごめんなさい、日本のパソコンと英語のパソコンでキーボードの配列が違うので…。アットマークは見つからないなあ。じゃあ、いいです。

ちょっと別の追跡調査があるのですが、残念ながらお見せできないようです。その内容は、今回の経年分析で出た結果を元に、どの分野を重点課題と設定して改善努力へ結びつけるかという教員へのアンケート調査です。今回の経年分析で学生の回答が思わしくなかった項目が幾つかあります。まずは学習経験としてプレゼンテーションをやっている学生が減少傾向にある。この原因は、おそらくオンライン授業の増加にあると思われます。例えば、同じ科目だとしても、2007年時点では教室の対面授業で口頭発表していたのが、2013年でオンライン授業に鞍替えした際に口頭発表の課題が外されたであろうことは容易に想像できます。もう一つは、数字の読解能力に関する学習経験の数値が減少傾向にありました。

今、お見せしようと思ってたのは、こうした改善点を浮き彫りにした後の、改善努力への移行に関する調査です。例えばですが、教学担当筆頭副学長が「今回の経年分析で改善すべき点は4つあるから、教員は努力するように」とお達しを出したとしましょう。それだと、教員は「何だ、上からの押し付けか」とか「執行部が学問の自治を侵害している」と声高に反発するのが目に見えています。このような状況を回避する為に、私なりに教員がどの項目を優先に改善するかという意見調査のアンケートを実施しました。これなら、上からのお仕着せではなく、教員が自分達で学習重点項目を決定したという形になりますから。

まず、アンケートの内容は簡単で、「今回の経年分析で4項目の改善すべき点が出ました。先

生方、来年の学習目標としてどれか2つ選びませんか。その2つを目標に掲げて、全学科の先生たちで授業法改善について議論してみませんか」という形で質問を5つぐらい設けました。一般的に教員アンケートは、先生方が忙しいのを言い訳に回答率が思わしくありません。なので、質問項目は物凄く短くしてあります。

今月でアンケートの回収が終わるので、予定では1月に春学期が始まったときに結果を教員に公表しようと考えています。アンケートの質問では、単に「改善点を選んでください」という、取り組み課題の優先順位だけではなく、「先生方の中で、ここで挙げられている改善点に対して既に取り組んでる事例があれば教えてください」と自由記述欄を設けてあります。これを呼び水にして、どこ学科の誰々先生はこういう授業法を展開しているとか、こんな教材を使っているという情報を教員の間で共有していこうと考えています。そうすれば、上からの押し付けというよりは、草の根でボトム・アップで改善活動を推進するのだという雰囲気をつくらせていこうと思っています。学生アンケートについては、こんな内容でよろしいですか。

ご質問者：そのプレゼンをするときに、例えばNSSEだと全体データとか、あとベンチマークとかの指標が返ってくると思うんですけども、そういったものを合わせながら示すといったことはしないんでしょうかね。

本田：実は、それが出来なかったんですね。その理由は、2010年と2007年の調査で競合校を選んだ条件と、2013年で競合校を選ぶ条件が変わったからです。勤務校は歴史的にコミュニティー・カレッジで始まったのですが、ここ数年で四年制大学に改組したことで、学部構成や教員資格も大幅に変わりました。そこで、競合校を選ぶ基準も変わったのです。なので、NSSEの本部が提供する競合校のベンチマークにはあまり依存しなくなかったわけです。

もう1点は、2013年でNSSEの質問項目が変わりました。だから、いくら競合校が同じでも、経年分析でのベンチマークができない部分が多くあります。

競合校との比較には問題がありますが、一応は参照しました。例えば、学内データだけ見て、統計的にこれは思わしくないと突き止めた分析結果を、2次資料として競合校とのベンチマークでどのような傾向が出ているのかも確認した上で提示しました。うまく答えになっていますか。

岡田：質問よろしいでしょうか。私、ベネッセ教育研究所の者なんですけれど、ネッシーのような学生の学習実態調査と客観的な批判的思考力のテストと適性検査を併せたパッケージを大学さんに提供して、その分析を戻す仕事を以前していたんですね。そのときにいつも問題になるのが、統計的な有意差検定をどこまでやるかっていうところで、例えば100人しかいない学科のコースで、90人受けましたとなったら、もう全数調査に近いものなんですよね。でも、そこで検定をかけてしまうと、結構な差が出てくるのに有意差なしとなってしまって、そうすると、何も言えなくなってしまう。ほとんど悉皆で受けているものは、検定かける必要があるのかっていつも議論になります。ほとんど全員対象の、ほぼ全員に受けてもらっているような調査で検定をかけることに意味があるのかなと。

本田：でも、検定をかけるといっても、どことどのグループを比べるのかによりますよね。例えば前年の学生調査アンケートと今年の学生調査アンケートだとか、ある特定の学科と全

学の比較するのか。今、私の方でご質問の意味が完全に理解できているかわかりませんが、私の分析の場合には全体数と標本数の問題は、どのグループとどのグループを比較しようとしているかが問題になると思うのですが。違いますかね。

岡田：例えば比較の対象が全国のサンプリング調査であればいいと思うんですけど、その対象が例えば前年と今年の差だったり、A学科とB学科の比較で、比較対象になっているグループは、母集団の大部分の人が回答している調査になるので、検定ではなくて、信頼区間みたいな形で出したほうがいいんじゃないかなって思うんですけど、とにかく全部、検定をかけてくれみたいな話になってしまうんですね。

本田：ごめんなさいね。私は統計の専門家ではないので、なさってることが十分に理解できているかわからないので、臆測を通じてお答えします。私の普段の分析では全数と標本はあまり大きな問題として扱わないようにしています。というのも、それはGeneralizability（データ結果の一般化）の話なので、それはサンプル数が多ければ多い程、良いわけですよ。

具体的な事例だと、建築学科のプログラム適格認定に提出するデータ分析があります。先にご紹介したNSSEという学生調査を建築学科の学生と全学の学生で比べてみたら面白いのかなと考えました。建築学科でアンケートに回答した学生数は、プログラム内では回答率が高いのですが、学生収容人数が少ないので、全学生の回答数に比べればかなり少ないわけです。この2つのグループをt検定に掛けるのはサンプル数の違いから無理がありそうだと思います、ベミジ州立大学の藤原さんに相談しました。彼の回答として「スタンダード・エラーの値があまりにも違い過ぎるから、それはやめたほうがいいんじゃないですか。私なら記述統計だけにしておきますけど」と助言を受けました。また、そもそもアンケート調査の回答率の低さや自己申告のデータの信頼性の疑問があるところに、無理に検定を掛けなくてもという話になりました。おそらく、全学生の回答に更にサンプルを掛けて、建築学科のそれと同程度に落として検定を掛けるのはできるのでしょうか、そもそも小さいサンプル数にそこまで労力を掛けて分析したところで、どれだけ結果を活用してもらえるかわからないので止めにしました。

岡田：基本的にはこれは抽出調査の結果なので、検定も全部かけるという扱いですかね。

本田：そうですね。あと、学内でもサンプル数でよく聞かれるのですが、学生調査だったら、30パーセントから35パーセント取れるのが平均的です。専門職の集会などで他大学の人に聞くと、大体どこもそういう回答率ですし、そこに達してればいいのかなと思うんですね。

ただ、それはあくまでもIRやアセスメントの専門家の経験値から出た数値なので、例えば社会学で批判的な教員だと、私のデータ分析を大学全体に一般化できないと噛み付いてきます。まあ、これはよくあることなので、私なりには「確かにそうかもしれませんが、それで今回出た分析結果を全て無視するのでしょうか、3割の学生でもこういう結果出てるけど、それを参考にして改善につなげることはできませんか」と切り替えます。正直、サンプル数で挑戦してくる教員は性格が悪かったり、もともと改善する気がなく、みんなの前で私を批判して、いい気になりたい人もいます。そういう批判的な教員の存在は全学で知れ渡っていますから、他の人たちも「ああ、あいつ単に批判したいだけなんだな」と見ているだけのようです。

多少、暴露話というか愚痴めいた話になりましたが、正式には「手法の限界があります」と発表ではあらかじめ伝えておきます。「標本はこれしか取れてないし、危なっかしい統計の検定なんかもかけてないです」というふうに言って、「手法の限界に留意して、分析結果の解釈には気を付けましょう」だとか、質問を受けても、「そこは答え難いです」と正直に分かる範囲だけを答えるようにしています。

補足：当日は質問の用語と趣旨を十分に理解できていませんでした（例えば、Confidence Intervalが日本語で信頼区間と呼ばれるなど）。私には全数調査を扱うことが稀なので、妥当な分析手法を検討する知識が欠けています。抽出調査の追記としては、全数と標本であまりにも違いのある変数（例えば性別、年齢など）があれば、Weight（重み付け）を掛けて分析しています。

岡田：ありがとうございます。

本田：では、2時40分になりましたが、まだ質疑応答を続けましょうか？ いいですか。

山本：ここまで本田先生からは、いくつかの具体的なトピックを取り上げながらアセスメントの本質までご報告いただきました。皆さんからも、いろいろとご質問もあるかと思しますので、短い時間ではありますが、質疑応答の時間とさせていただきたいと思います。本田先生、今のお話の中、もしくはお話の外でも、ご関心のある点があれば、というところでご質問を受けてくださいますか。

本田：そうですね。

山本：また、本日の資料で、触れてない部分も若干あるかと思しますので、そのあたりからご関心のある点もありましたら、ご発言いただければと思います。

本田：例えば、おわりにの部分も何も言ってないですし。

山本：この部分は、実務を担当している者にとっては、ちょっと気になる所ですね。

本田：愚痴ですからね、「おわりに」は。

山本：では、ご質問、お願いします。

佐藤：一ついいですかね。スライドの3ページは、僕の中では一番面白かった所だったんですけど、この学習成果とは何ですかというときに、アメリカの場合、左側ですよ。日本は右側という話はすごい面白くて、批判的思考力が高いからといって就職率に結び付かないみたいな議論がありますよね。

本田：批判し過ぎて就職できないとかですね（笑）。

佐藤：でも、1つ気になるのは、日本がなんで右を気にするかというと、特に私立の場合ですと、それがイコール大学の教育を社会的に説明するときに分かりやすいわけじゃないですか。批判的思考力高まりましたよ、コミュニケーション高まりましたよと説明するだけだと、分かりづらいというのが正直あるんですね。

本田：分かりづらいですね。

佐藤：アメリカでは、そういう議論はないのかなと。例えば、日本はすぐエンプロイアビリティ一みたいな議論じゃないですけど、雇用できるような人材を育成しろというような要請が

強いから、どうしても右側の所がこういうふうになって、割り切つてできないんですよ、大学の教員としては。だから、そこはアメリカがうらやましいなと思って。教員はこっちだけやっていたらいいんだって議論はうらやましいなと。アメリカでは、そういう議論は、どの程度まであるのですか。むしろないんだつたらいいんですけども。

本田：いえ。実は2つ例があります。1つはフォー・プロフィットと呼ばれる営利大学。カプレンなど大学入学試験の参考書や授業を開催している予備校めいた学校は日本でも知られますか？ 営利大学のお金の稼ぎ方は、失礼ですが、社会の貧困層にかなり近いところの人達に「学士課程を取れば就職もできて、次のソーシャル・ステータスにいけますよ」みたいな誘惑をして募集を掛けています。それで、入学手続きの段階から多額の奨学金をエサにして入学させます。ところが、奨学金による学費割引目当てで安易に入学した学生の多くは、卒業できないわけです。アメリカでは当該大学が出す奨学金の他に、連邦から出る奨学金があります。また、学生は複数の奨学金と学生ローンを組んで学費を賄うことが多いです。営利大学がどうやって儲けているかという、まずは入学の勧誘として自校の奨学金を初期投資として使い、入学後は学生が持ってくる連邦の奨学金で学費を吸い上げるという形で成り立っているわけですね。それで、多くの学生が卒業できずに負債だけ抱えます。

これではイカンということになって、2、3年前に連邦政府の監査が入って、そうした大学には学生募集に不正が無いかの検査や就職率のデータを出すよう要請しています。ただし、この話は全米の高等教育の中では例外的な扱いだと思います。実際、営利大学の学生数は全体で見たらかなり少ないというデータが出ていますし。

営利大学のような乱暴な話は、州立や非営利の私立でも上の層の大学には当てはまりません。例えば、ハーバードなど、アイビーリーグの大学で「就職率出してください」と言っても、親がみんなコネ持ってるから、就職率100パーセントでしょう。または、卒業後に自分探しで1、2年遊んでいるから就職率が低くなる場合も考えられますが、彼らはどうせ修士課程に進むことになります。別の例で、私の勤務校のような州立の中規模大学だと、文理学部というよりは看護や歯学衛生学科など実学系の学科長はアンケートで就職率を調査していますが、その程度です。

あと、もう1つの事例はソーシャル・セキュリティー番号を使った雇用データの確立です。近い話だと、日本ではマイナンバー制が導入されますよね。アメリカは既にソーシャル・セキュリティー番号があるので、そのデータを使って労働省との協同プロジェクトで、メイン州内にいる州立大学の卒業生の就職先が分かるようになりつつあります。私は試行データしか見たことがありません。データ構築はかなり出来上がりつつあるようですが、データのヒモ付けが2、3年度分くらいしか進んでいなかったようです。ただ、このシステムの限界は、卒業生で州外に引越した人達の就職データが追えないことです。こうした取り組みは、私の限られた知識ですが、おそらくメイン州が進んでるはずですよ。マサチューセッツ州やニューヨーク州では、同じことをやろうと思っても、州政府の図体が大きすぎたり、官僚的な手続きが多くあるので、なかなか立ち上げられないと思います。ただ、他の州で力を入れているのが、小学校から大学までの個人のデータをつなげようというシステムを構築しています。ただ、やはり大学卒業以降はやっぱり難しいので、そこま

でできてるの州は少ないと思います。

浅野：一つは補足になるかは分からないんですけど、つい3年ぐらい前のAIRでよく“Gainful employment”への言及が多かったと思うんです。

本田：それです。

浅野：意味としては、アメリカの場合、学生さんがローンを抱えて勉強する、つまり自分で教育に投資するわけですね。そのため、アメリカの大学が問われているのは、学位、その大学で教育受けたことによって、その投資に見合った収入が得られるかっていうのが問われている団と思います。特に、オバマ政権は、強く求めています。だから、就職率というよりは、大学に行って学士を取ったことによって、大学に行かなかった層の学生とどれだけ収入差が生まれてるかっていう、そこのデータを見せろっていう形で対応を迫られていると思うんですけど。

本田：まさにおっしゃる通りです。おそらく、教育社会学系の研究分野でそうした調査があるかと思っています。その辺は全国規模の調査でしょう。ただ、個別の大学で、そういうことをきちんとやってるかって言われると、そこまで手が回らない。

浅野：そうですね。それは、やっぱりデータ取りに行くってことですよね。ありがとうございます。

本田：いいえ。

山本：今、このスライドがありますので、このスライドの関係の所でご質問のある方はいらっしゃいますか？私からもこのスライド関係でよろしいですか。学習成果のお話で、左側の「コミュニケーション」、「批判的能力」は、獲得する能力になるわけですね。本日の先生のお話の中では、日本でいう科目の話、コースごとの学習成果のアセスメントのお話と、ア krediteーションなどの対応からのプログラム単位の学習成果のアセスメントのお話と両方あったと思います。大学としてどちらを重視するかですとか、もしくは本田先生のお仕事の範囲として、どちらの比重が高いのかなど、それぞれのアセスメントの目的やアセスメント専門家の関与方法や役割の違いについて、コメントをいただければと思います。

本田：今のお答えはこのスライドなのですが（題名：評価と診断の違い、本稿14頁）、左の行を見ていただくと、プログラムの次元なのか、科目の次元なのかで、科目のレベルはもう個人の先生のアカデミック・フリーダムに任せてあるので、そこに私が踏み込んでいくことはないですね。今、適格認定（Accreditation）でも厳しく言われてるのが、プログラムの次元で教員が集団としてカリキュラムの点検や授業法の改善に取り組んでるかという、「集団として」というのがキーワードなんです。正直、皆さんやってないですから。その集団的に学習成果を測定し、改善につなげるお膳立てをするのが、私みたいなアセスメント専門職の仕事です。教員が集団で取り組むという土壌さえできれば、あとは学習成果の活動の中での気付きだとか、反省点が生まれてくるのだと思います。そうした教員の集団による議論があったうえで、今度は各教員が自分の授業だとか、教えるときの科目の次元に戻って改善してもらおうという、そこは循環的なつながりなんですよ。お答えになってますか。

本田：他になければ、もうおわりにの所でまとめましょうか、50分か。

山本：そうですね。

本田：「終わりに」という部分を何も話していないので、そこを簡単に済ませるか、またはもう参加者アンケートに移るかです。

浅野：時間の関係もありますので、そろそろ最後のまとめに。

おわりに

学外の環境

- 伝統的IRの行方(Data Warehouse/BI)
- 外部コンサルティング会社(EM/Advising)
- MOOC
- 不公平な大学間競争

補足：このスライドの説明は飛ばしました。

おわりに

学内での働き方

- 非定型業務を動かせる
- 教員を支え、頼りにされ、時に気付きを与える
- 執行部には選択肢を提示する
- 基本的には黒子

本田：どこら辺が面白いのかな。『専門家とは?』という話を手短かにまとめてみます。

定型業務ばかりやってる間は、専門家とは言えないのかなと思います。アメリカのIRとアセスメントの比較で考えて見ましょう。まず、IRはデータベースからデータを引き落とし、主に執行部に報告書を提出し、残りは学部学科長からの単発的なデータの依頼に

応えていけば事がすみます。対照的に、私が携わるアセスメントの仕事は教員相手に測定の設定を構築して、学生の論文を査読したり、統一試験の実施でデータを集めてきます。皆さんも教員と仕事をして感じているかもしれませんが、教員は概ね個人主義で集団として動くことが出来ない。その中で舵取りをするのだから、時間的にも、精神的にもかなりの重労働です。一般的に、アセスメントの仕事はストレスが溜まり、辞めていく人も沢山います。また、私の場合は教学担当の筆頭副学長に付いている事もあり、新しい仕事が次々と降りてくるので、2週間後に自分が何やってるかなんて読めないですから。

IRに戻りますが、教員と比べて、執行部は経営管理という意味では合理的な集団なので、データの理解と活用の面で教員相手よりもかなり楽です。私はIRの専門集会でも研究発表をするので、IR関係者との交流があります。そこでは、「本田はアセスメントなんだ。大変だね」とか、「俺はアセスメント、絶対行きたくないから」みたいな本音を非公式に告白してくれる人もいますね。

更に、学習成果の仕事に就くからには、やっぱり博士号が必要です。私は修士号を取ってから、日本の私立の職員として始めました。そのとき、自分は物知りだと勘違いしてたんですが、アメリカの博士に進んで考えを改めました。

私自身がアメリカの博士課程に進んだ経験だけでなく、日本からの調査依頼でアメリカの大学院入試について面接調査を実施したこともあります。ある研究科長に聞いた話ですが、「修士と学士の差はあまり無い」そうです。というのも、授業を何個か多く取ればいいだけだから。ところが、博士と修士には天と地の差がある。だって、修士の卒業率は多分8割は優に超えていると思いますけど、Ph.D.は私の知ってるのは、90年代中盤の全米調査で少し古いですが、3割くらいの人しか取れない。そこを通り抜けてるか、抜けてないかで、教員の見目が正直違います。だからといって、「日本の評価担当者がみんな博士課程取らなければ駄目だよ」と言うつもりはありません。ただ、アメリカで戦うんだったら、特に学習成果であれば博士号は必須です。

2つ目の、教員を支えるのはもちろんで、気付きを与えること。時には、わざと仕事の詰めを自分で全て片付けてしまわずに、教員に「どうしましょうか」と議論や意思決定に参画してもら

うことです。そうしないと、僕が全部学習成果の仕事やっていて、先生たちの外部委託（アウト・ソーシング）みたいになっても困るので。教員達が自分達で決めて、運営しているプロジェクトだという意識を持ってもらえるような段取りを心がけています。

3つ目の執行部には選択肢を提示するというのは、私にとって大きな教訓でした。日本でも、多分3年とか、5年前にアメリカのIRの人は改革推進者か、学内コンサルタントみたいなイメージがあったかと思います。でも、私はその考え方に完全に合意しているわけではありません。これはあくまでも私個人の推察ですが、データ持っていると、自分はまだ改善案を既に知ってるような錯覚に陥るようで、学部長会議、副学長会議、学長会議とか、上の所で改善案を提言する場合があります。そうすると、「それはおまえの仕事じゃないだろう」と他部署の上職者から反感を抱かれます。IRは入学関係のデータを分析しているとしても、学生募集や在籍管理の仕事はEM部長の担当ですし、在籍に関する財務分析は財務担当副学長がいて、そこが予算配分の最終判断をしますから。

アメリカで結構シビアなのは、組織図で職務権限が決まっていますから、自分の権限以上の意思決定をすると、周りからすごい噛み付かれることがあります。私自身の失敗談ですが、「本田、おまえの職務権限以上のことを、俺に押し付けてないか」とか言ってくる人が居るんですね。権威主義で性格が捻くれていてと言っては何ですが、そういう人ほど、こうした権限だとか誰が誰に命令したとか、事細かに批判してきます。

だから、私が普段から実践しているのは常に選択肢だけを与える。選択肢を与えるときにも、データの分析では3つぐらいまで選択肢が考えられますが、実際の現場で働いている教職員の経験も質的な情報として考慮すべきですといます。また、私は教員評議会や、教学担当副学長会議にも出ているので、そこで得た自分の観察も意思決定の議論の際に盛り込みます。例えば、この3つの選択肢を考える上で学内政治ではこういうところを配慮した方が無難ではないかという、考える材料までお膳立てしてあげて、あとは腰を低くして上に決めてもらう。

もちろん、上職者や教員でも器量のある人は、下からの意見に耳を傾けます。その信頼関係を基に、データ分析から見えてくる提案を聞かせて欲しいと言われるときもあります。その際は、遠慮なく自分の考えとして、現段階の意思決定とそれが将来の方向性にどのような影響を及ぼすであろうか、また自分ならどのように物事を進めるかなどを伝えます。

それが無いままで上職者会議に出て、突然、関係部署のディレクターとの根回しもせずに問題を指摘し、改善案をぶち撒けると嫌われます。これは個人の性格と哲学にも寄るのでしょうが、私はIRやアセスメントは黒子だと思っています。それが、自分はデータを持って改革者ぶって前面に出たい人は、敵を作れど、信頼は得られない。いやあー、だって、みんないいかっこしたいですもん。アメリカでは特にね。このような状況下だからこそ、黒子でいいっていう人の方がIRや学習成果の仕事には向いてると思います。

格言

データがないと始まらない

けど、

データがあっても始まらない。

浅野：最後の格言。せっかくだから、ご披露いただいたらよいのでは。

本田：格言？ああ、これはいつもこんな状況ではないのですが、本質的に感じていることです。データがないと始まらないけど、データがあっても始まらない。

データがないと始まらないと言ってくれる上司がいるのは、すごいありがたいことです。彼は専門がコンピューター・サイエンスな

31

ので、データベースも知ってるし、データの使い方をよく知ってるんですね。また、案件によつてですが、上司からはデータを見てから意思決定したいと言われるときもあります。また、最近だと英語学科の教員の学習成果の議論では、個人的な学生観察や教育哲学が飛び交うので、「まずはデータを取って議論しよう」という先生が1, 2人出てきました。これは本当にありがたいことです。

データがあっても始まらないのは教員評議会が典型かなあ。多分、政策立案の段階になると、過去の数値データだけじゃ話が進まないですから。あるデータ分析の結果だけで、今までの慣行を突然変えるのって無理な話ですよ。また、データには過去の検証と将来の展開の大きく2つがありますが、後者は単純に意見の違いや思い描く将来像の違いが大きな割合を占めるでしょう。そうなると、意思決定におけるデータの重要性は相対的に低くなる場合があります。

ご清聴いただき、誠にありがとうございました。

<以下説明に至らなかったスライドを掲載しています>

ルーブリック

個別科目での活用

- 教員の頭の中にある基準を明確に外部化する
- 学生からの質問や苦情が減る
http://www.niad.ac.jp/n_kenkyukai/1229401_1207.html
(日本語の参照例)

教育課程での活用(ENG101)

- 教員間の共通理解を拡大する
- 1つの小論文を2名の教員が査読
- 査読結果が分かれた場合、2名の教員が結果を議論するか、3人目の教員を査読に当てる

16

学生による授業評価

- 学内の文脈
- 質問内容: 授業方法と学習成果
- 質問項目と教員組合
- 回答率と同僚評価

質問例

<https://teaching.berkeley.edu/course-evaluations-question-bank>

学外の業者

<http://ideaedu.org/services/student-ratings-of-instruction/>

17

情報リテラシー共通試験

- 情報リテラシー(SAILS)
<https://www.projectsails.org/>
- プロジェクト管理
- 学内でのデータ発表
- 結果をどう活用するか？
 - 教員と図書館員の協同

18

ポートフォリオ

- 総合的に学生の習熟度をみれるが、準備と作業が大変
- 学習成果
<http://www.tdupress.jp/books/isbn978-4-501-62740-9.html> (日本語の参照例)
<http://as.wiley.com/WileyCDA/WileyTitle/productCd-0470503769.html>
- 教員審査
http://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/28803/1/11_P129-140.pdf (日本語の参照例)

19

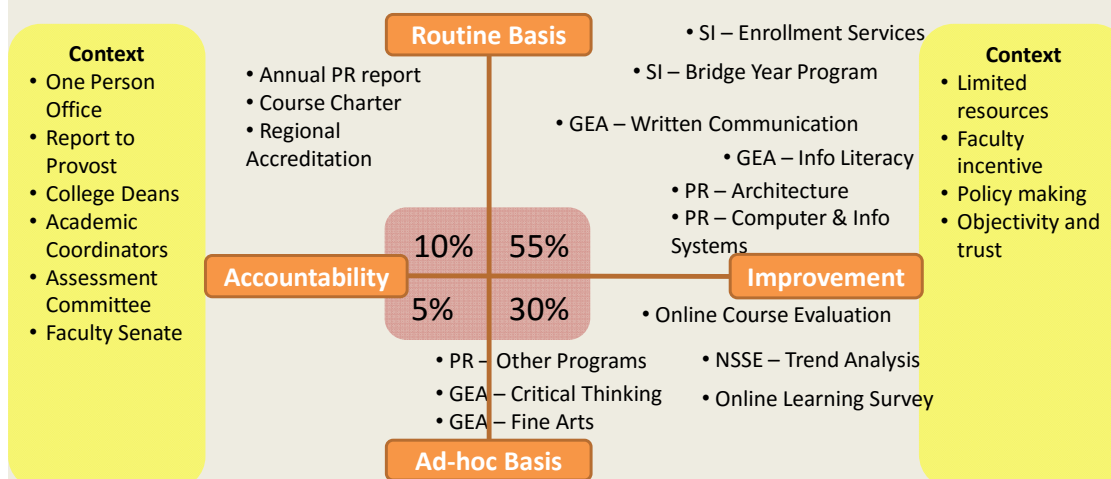
付録: その他業務

- 戦略計画の診断
- コロンビア大学College Educational Quality Project
<http://collegeedquality.weebly.com/>
- 高大連携プログラムの診断
<http://www.nacep.org/>

20

四象限による業務分析

KEY:
GEA=Gen Ed Assessment
PR=Program Review
SI=Strategic Initiative



データ活用の複雑な様相

• データの

- 収集
- 分析
- 共有
- 活用



23

IRとAssessmentの違い

IR

学生データ

- 入学
- 奨学金
- 在籍率
- 成績
- 卒業率
- 就職率

Assessment

アンケート調査

(間接指標)

- 新入生
- 学習実態
- 満足度
- 卒業生

学習成果データ

(直接指標)

- ルーブリック
- ポートフォリオ
- 試験
- 論文

24

