

# 学問基盤力テストの 分析結果

安田 淳一郎(山形大学 学術研究院)



皆さま、おはようございます。山形大の安田淳一郎と申します。私のほうからは、学問基盤力テストの分析結果ということでお話をいたします。私のお話大体 40 分ぐらいあるんですけども、前半はですね、この学問基盤力テストの目的やその開発の概要についてお話しします。後半は、分析結果についてお話しするんですが、若干ですね、テクニカルなところもありますの

で、できるだけ分かりやすくゆっくりとお話をさせていただきたいと、そういったことを心がけながらお話しさせていただきたいというふうに思います。

## 基盤力テストの目的

- 学生の学習達成度を**直接評価**し、**基盤教育の効果を検証**すること

山形大学の学士課程教育

4 年次

基盤力テスト (3 年次)

2 ~ 3 年次  
基盤専門教育

基盤力テスト (2 年始業時)

1 年次  
基盤共通教育

基盤力テスト (入学時)

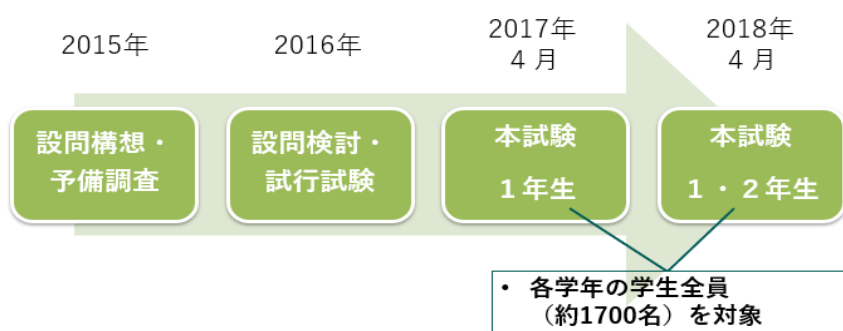


11

まず学問基盤力テストを含む基盤力テストの概要からなんですけれども、先ほど千代さんのほうからありましたが、まずこの基盤力テスト何のためにやっているのかということを確認させていただきたいと思います。基盤力テストの目的というのは、学生の学修達成度を直接評価し、基盤教育の効果を検証することとさせていただいております。こちらが本学の学士課程

教育を表したものなんですけれども、1 年次に一般的な教養教育にあたる基盤共通教育というものを行ってまして、2 年次に基盤専門教育ですね。そして 4 年次と続いていくわけなんですけども、基盤力テストをこの教育の各段階において実施しております。入学時に 1 回目、2 年始業時に 2 回目、そして 3 年次に今後実施していく予定です。このように教育の各段階において、学生さんの学修達成度を可視化しようということがこの基盤力テストの目的であります。

## 開発スケジュール



次にこの開発のスケジュールについてお話します。開発するスケジュールなんですけれども、2015年にこのテストの構想が始まりました。その後、設問開発及び予備調査を繰り返してまいりまして、2017年ですね、昨年の4月に本試験の1回目、今年度4月に本試験の2回目を実施したところであります。本試験については、各学年の学生さん全員を対象しております。各学年、1700人おり

ますが、全員を対象にしてこのテストを実施しているということになります。

## 開発の概要

### 開発体制

- 基盤力テストWG(基盤教育企画部の各専門分野の教員4名)

### 分野

- 数的文章理解, 数学, 物理, 化学, 生物
- 各分野で, 30~45問程度を作問

### 開発方針

- 知識の記憶だけではなく, **概念を理解**していることを測定できるテストを開発する

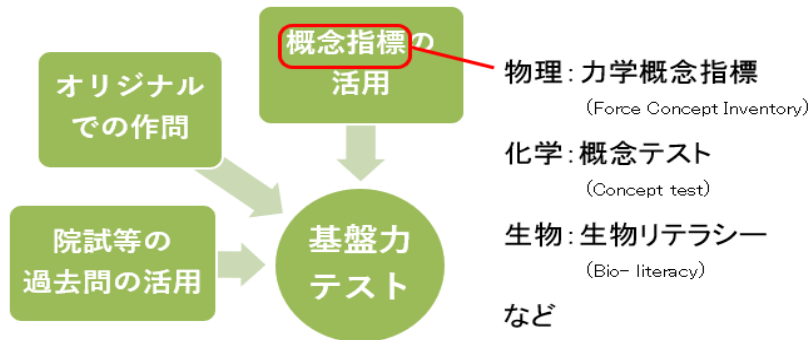
この基盤力テスト、開発の概要なんですけれども、まずこの開発の体制といたしまして、学内で基盤力テストワーキングというものを組織しております。これはですね、旧教養部に対応する、本学の基盤教育企画部というところがありまして、この各専門分野基本4名でこのワーキングは構成されております。学問基盤力テストの分野は5分野ございまして、数的文章理解と

数学・物理・化学・生物と。

数的文章理解と申しますのは、先ほど千代さんのほうからもありましたが、統計学の初歩のような科目で、データの解析であったりとか、データの分析にかかわる科目です。当初の段階で、この各分野で30から45問程度作問いたしました。そのときの開発の方針なんですけれども、知識の記憶だけではなく、概念を理解していることを測定できるテストを開発するというので、例えば化学の場合ですと、ある化学式覚えているかどうかとか、あとは生物の例で言いますと、ある生物の器官の名前を覚えているかどうかとか、そういった記憶を尋ねるのではなくて、覚えているかどうかということではなくて、各科目の概念をきちんと理解しているかどうか。ここまで測定できるようなテストを開発すると。こういった方針で作問を進めた次第です。

## 開発の手法

### 様々なリソースの活用

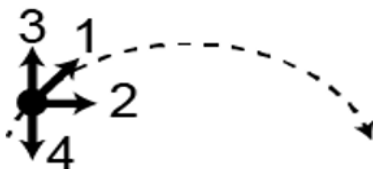


その作問の手法なんですけれども、1からテストの作問をしようとすると、非常に手間がかかってしまいますので、さまざまなリソースを活用しております。オリジナルでの作問はもちろんなんですけれども、院試等の過去問の活用であったりとか、あと概念指標と呼ばれるものを活用しております。この概念指標と申しますのは、1990年初頭ごろから米国で開発が進められてきているもの

で、このように各科目で、概念指標と呼ばれるものが開発されています。物理の場合ですと、力学概念指標、化学の場合だと概念テスト、生物の場合だと生物リテラシーですかね。こういった概念指標と呼ばれるものを参考にしながら、テストの作問の開発を、テスト作問をしているということです。1問だけ、作問の試作例をご紹介しますと思います。

### 設問試作例（物理）

問. バasketボールの選手がシュートした。破線のような軌跡をたどるとき、ボールが手を離れた少し後に下図の地点でボールに実際にはたらいっている力を選べ。



これは物理の設問の例なんですけれども、少し読み上げますと、「Basketボールの選手がシュートした。波線のような軌跡をたどるとき、ボールが手を離れた少しあとに、下の図の地点で、ボールに実際に働いている力を選べ」と。1・2・3・4と矢印がありまして、この中から1つ矢印を選ぶわけなんですけれども、このような形式になっています。基盤力テストのすべての

設問も、このような多肢選択式の設問になっております。ただ、こういうふうな多肢選択式の設問ですと、いわゆるまぐれ当たりというものが生じる可能性が、一定の確率であります。そういったまぐれ当たりのようなものが、高確率で生じているような設問がないかどうかと。その設問の妥当性ですね。これについては、きちんと調査の段階で調べております。

## 設問の妥当性

### 予備調査での分析手法

1. 記述式回答の挿入, 事後インタビュー
2. 項目反応曲線による分析

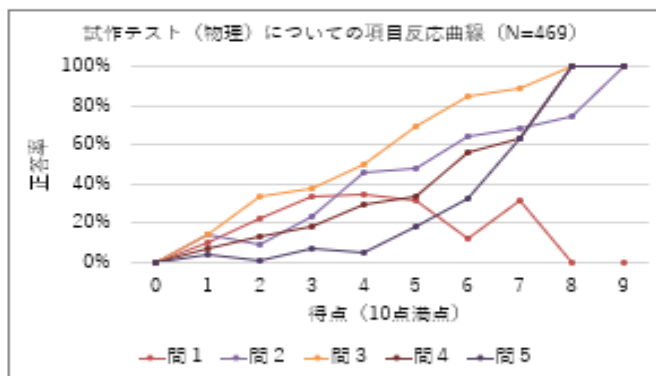


その分析の方法なんですけれども、大きく分けると2つあります。1つ目は記述式解答を挿入、事後インタビューということで、多肢選択式の設問のあとに、なぜその選択肢を選びましたかということを自由記述で書かせる。そういった調査票を作成しまして、それで調査を行ったりとかですね。いろいろ調査で、あと事後インタビューといたしまして、いろいろ調査に協力して

いただいた学生さんを、協力いただいて、詳しく解答時の思考過程を尋ねたりとか、そういったことをして、何かその設問のほうにおかしなことが起こっていないかということを確認しました。

## 設問の妥当性 (つづき)

### 項目反応曲線による分析例



もう1つの方法といたしましては、項目反応曲線による分析というものも行っております。これはですね、インタビューとかだとやはり手間が少しかかってしまいますので、より効率的にその妥当性を調べる方法なんですけれども、項目反応曲線、少しテクニカルな言葉なので、詳しく説明します。このグラフは、物理の試作テストについての項目反応曲線を表しています。こ

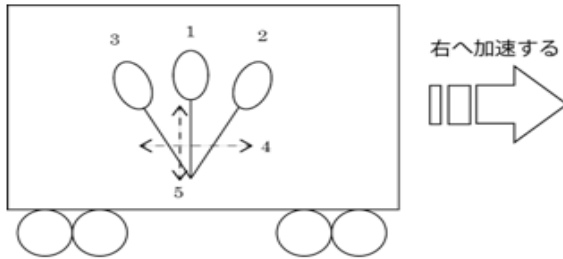
の試作テストに協力してくださった学生さんの数は469名でした。この試作テストは、10問からなるテストで、10点満点です。この横軸が、その試作テストの得点を表しております。0点から、ほんとは10点まで可能性としてはあるんですが、10点満点の学生さんはいなかったの、まあその10のところは見えてきません。縦軸は正答率ですね。各得点を獲得した学生さんの中で、それぞれの設問を正答した学生の割合はどれほどいるのかということ縦軸に表しています。例えばですね、このところに着目いたしますと、4点を獲得した学生さんの中で、この問3を正解した学生の割合はおおよそ70%であるということを示しています。この項目反応曲線なんですけれども、理想的には右肩上がりであってほしいわけです。物理は総合得点ですね。総合得点が高い学生ほど正答率が高い。これがまあリーズナブルな振る舞いなわけなんですけれども、これ1つですね、不思議な変わった振る舞いをしている設問が1つだけあります。これ問1ですね。はい。問1については、上がって下がるという、こういった

不思議な振る舞いをしておりまして、こういった設問については、やはりこれは妥当ではないといでこの調査分析の段階で外して除外しております。

### 設問の妥当性 (つづき)

#### 前スライドの問1

静止していた電車が一定加速度で右に加速し始めた。この電車内にある、浮いている風船は電車の中にある乗客から見てどのようになるか。



具体的などといった設問だったかというのは、その次のスライド、配布資料のほうにございますので、興味のある方は見ていただきたいと思います。

### テストの出題形式

- スマートフォンのアプリ「YU Portal」を使用



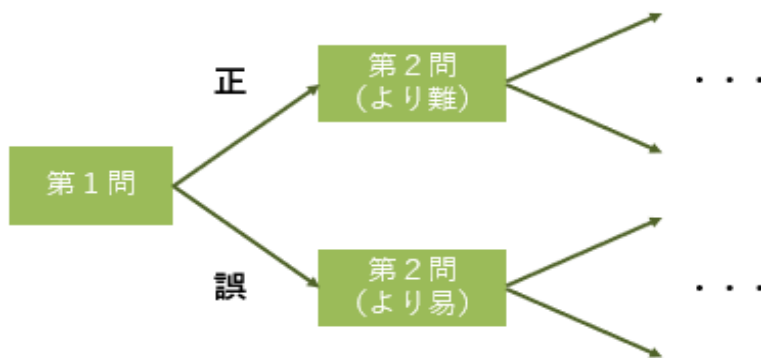
- コンピュータ適応型テストを採用
  - 項目反応理論(Item Response Theory, IRT)に基づき、受験者の解答に応じて出題

こういった形で開発した作問を用いて、基盤力テストを構成されているわけなんですけれども、この基盤力テストの最大の特徴と申し上げても過言ではないと思うんですが、それは先ほど千代さんのほうから説明ありましたが、このテストは、学生さんがスマートフォンを通じて解答するところです。本学で独自に開発いたしました「YU Portal」というものを学生さん

さんのスマホにインストールしてもらい、学生さんがアプリ経由でテストに解答します。

このスマートフォンを使用すると。コンピュータを使用するということを、メリットを最大限に生かすために、この基盤力テストはコンピュータ適応型テストと呼ばれるものを採用しています。これは項目反応理論と呼ばれるものに基づいて、受験者の解答に応じて出題される設問が変わってくるという、そういった形式です。

## コンピュータ適応型テストの概要



これ分かりやすく説明するために図を用意いたしました。例えば、第1問ですね。第1問に、正解した場合は第2問で、第1問よりも難しい設問が出題されると。逆に第1問を誤答した場合には、第2問でより優しい設問が出題されると。これによって能力値が非常に高い学生さんが非常に簡単な設問を解く手間が省ける。逆に能力値がそれほど高くない学生さんが著しく難し

い設問を解く手間が省けると。こういったことによって、テストの時間を短縮できると。さらには、その学生さんの能力値に応じた出題ができるので、より精度の高い測定ができるというふうに言われています。

## テストの出題形式

- スマートフォンのアプリ「YU Portal」を使用



- **コンピュータ適応型テスト**を採用
  - 項目反応理論に基づき、受験者の解答に応じて出題
- 出題数: 各分野**5問ずつ**

このコンピュータ適応型テストで行われている基盤力テストなのですが、出題率ですね、各分野5問ずつとしております。これ昨年のAPシンポでも、5問というのを出したら、けっこうですね、あまりよくない意味で驚かれる方多くて、「5問で本当に大丈夫なんですか」という質問も出たことがあるんですけども、少しこのなぜ5問なのかという、その根拠をお示しするのに、

分かりやすいスライドをご用意してきました。

## コンピュータ適応型テストの概要



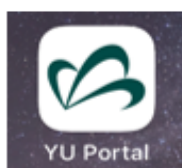
5問で、能力値を最大32段階にレベル分け可能

はい。これ5問、コンピュータ適応型テストの5問というのは、一般的なペーパー、紙形式の1問1点のテストの5問とは全く違うものなんです。1問1点の紙の形式のテストで、その受験者の能力を何段階に識別できるかということ、通常はその紙のテストだと、0・1・2・3・4・5なので、6段階ですね。6段階にしか分類、レベル分けすることができません。ただ、その

一方で、このコンピュータ適応型テストの5問の場合なんですけど、このテストでは、各設問に難易度が設定されています。このあとまたご説明しますが、難易度が設定されていて、どの設問に正答するかで、それを正解することによる得点能力が変わってきます。これによって、ある設問に正答するか、誤答するかの2通りですね。それで5回繰り返すので、2の5乗で32通りです。最終的な能力値というのは、理想的にはこの5問で32段階にレベル分けすることができるわけです。これは紙の1問1点のテストで言うと、31問分のテストに匹敵するということになります。こういった理由から、5問というとな少ないんじゃないかというふうに思われるかもしれませんが、これは全くそうではない、一般的にはその紙のテストでいうところの大体30問ぐらいに匹敵するんだということでご理解いただければというふうに思います。

## テストの出題形式

- スマートフォンアプリ「YU Portal」を使用



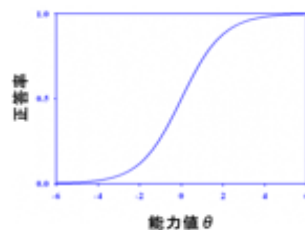
- コンピュータ適応型テストを採用
  - 項目反応理論に基づき、受験者の解答に応じて出題
- 出題数:各分野5問ずつ
- 設問毎の制限時間:3分
  - 試験時間は5科目で30分程度

続きまして、時間なんですけれども、設問ごとの制限時間は3分間というふうにいたしました。試験時間は5科目で30分程度を想定して開発いたしました。これは想定のとときは30分程度だったんですけども、実際にやってみたところ、どのぐらいになるかというのは、やってみないと分からなかったというところがありまして、その分析結果については、後ほどお示ししたいというふう

に思います。それで、分析結果のご説明のときに、どうしても必要になってしまいますので、項目反応理論について、最低限のことだけお話しさせていただきたいと思います。

## 項目反応理論 (Item Response Theory)

- 事前試験で各設問の難易度を推定



項目反応理論では、本試験の前に事前試験というものを行う必要があります。そこで各設問の難易度を推定します。この難易度というのは、こういった順番でその設問を出題するのとか、もしくはその受験者の能力値を推定するときに、どうしても必ず必要になるわけです。その難易度というのは、本試験の前の事前試験で行う必要かを、このデータを使って推定する必要があります。

## 開発スケジュール

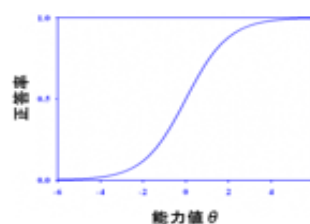


ちょっとスケジュールの話に戻るんですが、その事前試験もここで行っています。本試験の前年度2016年の12月ですね、1年生、文理混合1年生の集団、約60名を対象に試行試験を行いまして、そのデータを用いて各設問の内容を設定しております。推定された難易度を用いて、本試験で試験を行いまして、受験者の解答パターンに基づいて、その受験者の方の能力値と呼ばれるものを推定してい

ます。

## 項目反応理論 (Item Response Theory)

- 事前試験で各設問の難易度を推定
- 本試験で受験者の能力値( $\theta$ )を推定
- 本テストでは、1パラメータのラッシュモデルを採用



この能力値は、これ文字式です。シータと呼びますが、この能力値シータというものを推定しているんですけども、このシータというのは、その試験の得点のようなものです。ただ、従来のテストのように、100点満点とかそういったものではなくて、偏差値のようなもので表されています。偏差値の場合は、平均値が50



で、標準偏差が10になるように標準化されていますけれども、このコンピュータ適応型テストにおいては、この事前試験の被験者の能力値の平均を0、そして標準偏差値を1に標準化することが多いので、この基盤力テストでは、それに従って標準化を行っています。能力値の標準化については、また後ほど分析のところで詳しく補足の説明をさせていただきます。それで、このほかにもさまざまなモデルがあるんですけども、本テストでは、最もシンプルな1パラメーターのラッシュモデルと呼ばれるものを採用しています。

### 学生へのフィードバック


各科目の解答終了直後に、

- 4段階の能力レベル
- メッセージ



を**即時**にフィードバック

➤ **学習への動機づけ**を意図

残り: 2分29秒

  
**「ブロンズ」レベル**

物理学と身の回りの現象の関係について、様々な啓蒙書を読んだり、どうしてそのような現象が起きるのか考えてみたりして科学的な思考に親しみましょう。たとえば、虹の7色はどうしてそのような順番で並んでいるのでしょうか。そのほか、お湯を沸かすときに一番早く沸く方法、走っている電車の中でジャンプしたときに感じる力の種類など、みなさんが今いるんな専門分野で学んでいくときに理解しておく让世界が広がる内容がたくさんあります。



27

では続きまして、学生さんへのフィードバックですね。これは先ほど千代さんのほうからご説明ありましたが、少し補足説明をさせていただきます。これは今年度の4月から新しく始めた取り組みですね。各科目の解答終了直後に、4段階の能力レベルとメッセージを即時にフィードバックしております。4段階の能力レベルというのは、先ほどのご説明で、最大32段階レベル

分けできるというふうに申し上げましたが、それは理想的な場合の話で、実際にはもう少しですね、精度上がらなくて、4段階に識別すれば間違いはないと。そういった考えから、4段階のレベルに分けてフィードバックしています。このメッセージは、それぞれの科目及び能力レベルごとに異なるものを学生さんにお返ししています。これが具体的なメッセージの例なんですけれども、これは物理のブロンズレベルですね。能力レベルの名称なんですけれども、上から4段階あります。上からダイヤモンド・ゴールド・シルバー・ブロンズと、どこかのカード会社のステータスみたいな、そういった名称になっています。

なぜそういうふうな名称にしたかという、例えば「秀・優・良・可」とか、あとは「S・A・B・C」とか、何かそういった成績を暗示するような名称だと、少し学生さんのほうにちょっと理不尽さを与えてしまうところがあると。それなぜかと言うと、1年生の4月の段階で、理系の学生さんには、数学・物理とか、化学とか生物、すべての科目を受けてもらいます。ただ高校で、理科の選択で、選択しない科目についても、そのテストを受ける必要があるわけなんです。なので、高校で生物を選択していないにもかかわらず、1年次の4月の段階で生物のテストを受けさせられて、可とかCとか、そういった形の評価を聞くと、やはりその学生さんのモチベーションも下がってしまうだろうということで、こういったブロンズレベルとか、こういった名称にして、今後そのレベルを上げていきたいと思います、というふうにしていくと。そういった工夫があります。

このメッセージの内容なんですけれども、ちょっと読み上げますが、「物理学と身の回りの現象の関係について、さまざまな啓蒙書を読んだり、どうしてそのような現象が起きるのかを考えてみたりして、科学的な指標に親しみましょう」ということで、「あなたの物理の能力はまだまだです」とか、何かそう

いったん厳しい文言ではなくて、「今後物理について勉強していきましょう」「どういうふうにして勉強していけばいいのか」と、そういうアドバイスを与えるような学生さんの、学習の動機づけを高めることを意図した内容になっています。

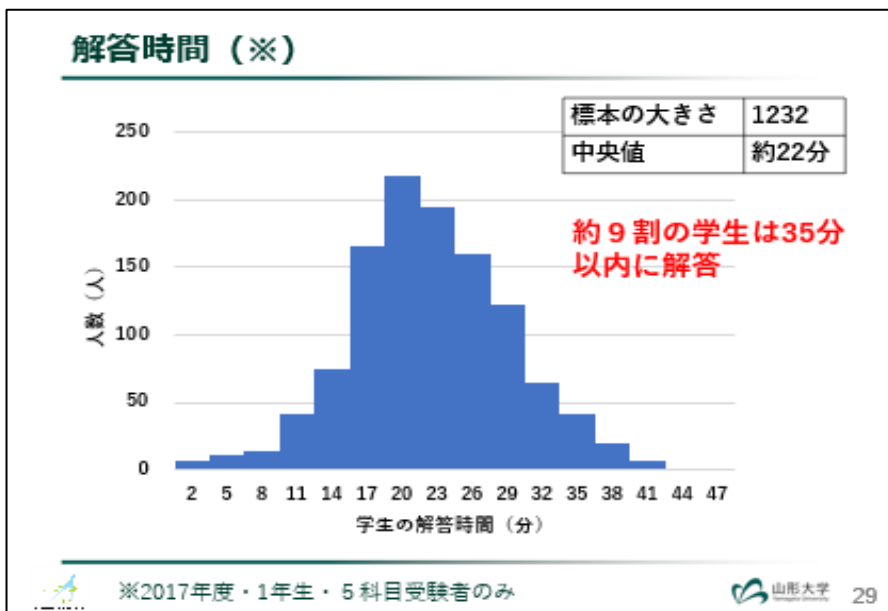
### 受験率

	数的文章理解	数学	物理	化学	生物
対象学部	全学部	理系のみ (文系は対象外)			
2017年度 1年生	99.3%	99.0%	99.0%	99.0%	99.3%
2018年度 1年生	99.2%	99.6%	99.6%	99.6%	99.5%
2018年度 2年生	89.2%	90.5%	90.2%	90.2%	90.6%

山形大学 28

それではここから、学問基盤力テストのデータ分析の結果について、ご紹介をしていきます。まず受験率からです。5科目ありまして、学部によって受験科目が異なります。数的文章理解については全学、数学・物理・化学・生物については理系の学部の学生さんのみが受験します。昨年度の1年生及び今年度の1年生については、ご覧のとおり受験率は99%を超えておりまして、

ほぼすべての学生さんに受験をしていただいております。今年度の2年生については、99%は超えていないんですけれども、それでもほぼ90%には達しているという状況です。この基盤力テスト、各学部・学科のオリエンテーションの一部の時間をお借りして実施しております。オリエンテーションに来ないとか、来てもテストを受けずに帰ってしまうとか、そういった学生さんがやっぱり2年生になると少し増えてしまうということで、10%ぐらい低下していますけれども、それでも9割は達成している状況です。



続きまして、解答時間の分析結果です。これは昨年度の1年生の5科目受験者のみのデータを取り出したものをヒストグラムで表しています。横軸が受験者の解答時間を分単位で表しておりまして、縦軸がその解答時間で解答した受験者の人数ですね。このようなきれいな正規分布のような形をしておりまして、中央値は約22人と算出されております。調べましたところ、約9割の学生さんは

35分以内に解答をしたということが分かりまして、これは当初30分程度で基盤力テストの解答を終えるというふうに想定していたんですけれども、おおよそその想定範囲内に収まっているかなと考えております。

## 分析結果（化学）：2017年度1年生VS2018年2年生

カリ キュ ラム	能力値 $\theta$ 1年	能力値 $\theta$ 2年	能力値 $\theta$ 差	$p$	$d$
全	-0.19	0.09	0.28	0.00	0.33
A	-0.14	0.31	0.45	0.00	0.50
B	0.38	0.44	0.06	0.55	0.08
C	-0.48	-0.56	-0.09	0.39	-0.13
D	-0.22	0.50	0.72	0.00	0.81
E	-0.05	0.32	0.37	0.00	0.46
F	-0.39	-0.43	-0.04	0.54	-0.06
G	-0.16	0.18	0.34	0.00	0.40
H	-0.36	-0.11	0.26	0.43	0.33
I	-0.45	0.16	0.62	0.00	0.72
J	-0.45	-0.37	0.08	0.22	0.13

✓ 16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		

続きまして、ここから一番重要な分析結果のご報告になります。昨年度の1年生と今年度の2年生の比較ですね。これはすなわち昨年度入学された学生さんが、昨年1年間、山形大学の教育を受けて、その能力をどれほど伸ばしたのか、もしくは下げたのか、もしくはあまり変わっていないのか。それを見るための最も重要な分析結果になります。分析結果の表はやや複雑な形になっていますの

で、丁寧に説明を差し上げたいと思います。説明を差し上げる順番なんですけど、化学から説明をさせていただきます。まずこの左側のところなんですけれども、これカリキュラムということで、AからJまであります。この「全」というのは、全ての学科・コースの平均を表しています。次に、この能力値シート1年とありますけれども、これは、化学の能力値の昨年度の1年生の4月の段階での能力の平均値を表しています。これ全なので、全学生の化学の能力値の平均値ですね。これはマイナス0.19と。能力値シート2年というのは、今年度の2年生ですね。彼らが2年生になって、この4月にテストを受けて、そのときの能力値の平均が0.09。その差が、プラス0.28ということで、およそ0.3ほど上昇したということになります。少し前のスライドで申し上げたんですが、このシートというのは、偏差値のようなもので、ある基準に基づいて標準化されています。その基準というのが、2016年度に行いました調査で、2016年度の1年生12月時点のシートの平均が0、シートの標準偏差が1になるように標準化されております。これより、ここにあります、シート1年がマイナス0.19というのは、1年生の12月時点が0なので、1年生の12月時点の能力値と比べれば、1年生4月時点の能力値は低いと。一方でシート2年0.09ということで、正の値になっていますけれども、これ1年生の12月と比べて2年生の4月は平均能力値が上がっていると。これはリーズナブルな結果であると言えると思います。次に、昨年4月から今年度の4月にかけて、能力値が0.28上昇したんですが、この差に意味があるのかどうか。偶然の可能性もありますので、それをきちんと見ております。これは統計学の基本的な方法なんですけれども、このp値というのを見ておまして、このpが0.05よりも小さい場合ですね。このときは、この差が有意であると。偶然ではなくて、意味のある差であるというふうに判定をしております。有意と判定されたところはこの緑色で色が塗られています。白色のところは有意ではない、偶然の差であるということなんです。

このp値に加えまして、このdと呼ばれるものも見ています。これなぜかと言いますと、このp値というのは、サンプルサイズが大きい、つまり受験者数が多い場合に有意になりやすいという問題点が指摘されておまして、その差が有意であるかどうかということだけではなくて、この差がどれだけ大きいのかということを見ています。このdというのは、「Cohen's d」と呼ばれるもので、日本語だと効果量と呼ばれるものです。このdの大きさによって、小・中・大、この3つに分類される

んですが、このシータが伸びた、増えたときにはオレンジ色で、減ったときには青色で表しています。この表を見ますと、最も顕著な結果というのは、このカリキュラムDですね。このカリキュラムについては、1年のときのシータの平均がマイナス0.2、2年生のシータの平均が0.50ということで、1年間で0.72も向上しています。この結果、当然有意で、効果量大ということで、このカリキュラムについては、その能力値が著しく上昇したということが言えるわけなんです。なぜこういうふうはこのカリキュラムDが大きく伸びているのかという、そこの解釈もしております、このカリキュラムDというのは、化学を専攻・専門とするカリキュラムなんです。なので、化学に関する授業をたくさん受けている学生さんが多い、化学を学ぶモチベーションも高いということで、これはまあ伸びて当然だよなというふうに思われるかもしれませんが、当然なんだけど、当然にできているところがやはり重要なわけです。化学の授業がたくさん開講されていたとしても、その授業で学生さんがみんな寝ている、試験は過去問を写すだけ、そういったタイプの授業を受けているのであれば、能力値はここまで上がらないはずなんです。

### 分析結果（生物）：2017年度1年生VS2018年2年生

カリキュラム	能力値 $\theta$ 1年	能力値 $\theta$ 2年	能力値 $\theta$ 差	$p$	$d$
全	-0.08	0.03	0.11	0.00	0.12
A	-0.13	0.12	0.25	0.00	0.27
B	0.42	0.71	0.29	0.00	0.32
C	0.61	1.00	0.39	0.00	0.42
D	0.09	0.21	0.12	0.24	0.17
E	-0.40	-0.38	0.02	0.79	0.02
F	-0.54	-0.43	0.11	0.59	0.09
G	0.00	0.01	0.01	0.91	0.01
H	-0.56	-0.60	-0.04	0.65	-0.06
I	-0.54	-0.52	0.02	0.94	0.01
J	-0.39	-0.19	0.20	0.36	0.34

✓ 16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		

では続きまして、生物ですね。生物については、先ほどの化学に比べると、このオレンジの部分には減るんですけども、カリキュラムA・B・Cにおいて、効果量小で能力値が伸びているということが分かりました。

### 分析結果（数学）：2017年度1年生VS2018年2年生

カリキュラム	能力値 $\theta$ 1年	能力値 $\theta$ 2年	能力値 $\theta$ 差	$p$	$d$
全	-0.10	-0.14	-0.05	0.12	-0.05
A	0.02	0.03	0.01	0.94	0.01
B	0.48	0.06	-0.42	0.00	-0.57
C	-0.96	-0.94	0.02	0.85	0.02
D	-0.07	-0.13	-0.06	0.57	-0.07
E	-0.25	-0.06	0.19	0.02	0.24
F	0.08	0.11	0.03	0.70	0.04
G	-0.01	0.14	0.15	0.04	0.19
H	-0.07	-0.27	-0.20	0.49	-0.23
I	0.03	-0.10	-0.13	0.48	-0.17
J	-0.48	-0.75	-0.27	0.00	-0.32

✓ 16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		

続きまして、数学です。これは5科目の中で最も良くなかった結果です。もう先ほどの結果に比べると、オレンジの部分が少なく、青いところが目立つわけなんですけれども、特にいまこのカリキュラムBについては、1年生のときのシータの能力値の平均が0.48で、2年のシータの平均が0.06、その差がマイナス0.42となっています。この差は有意で効果量が中ということで、

このカリキュラムBについては、もう数学の能力値の平均が著しく下がったということが分かつ

たわけです。この理由なんですけれども、カリキュラムBというのは、数学を専門とするカリキュラムではございません。さらに見ていただきますと、このカリキュラムBについては、この能力値のシータの1年生ときの平均が 0.48 ということで、これほかのカリキュラムに比べると、非常に高いんですね。入学してきた段階で、このカリキュラムに所属する学生さんたちは、すごい高い数学の能力値を有している。シータ2年生を見ますと、0.06 ということで、これは低いのかと言われると実はそうではなくて、ほかのカリキュラムの学生さんと同程度か、もしくは高いぐらいです。ということで、このカリキュラムBについては、この1年間で数学を学ぶ学生が多くなかったことで、入学ときに有していた数学の高い能力値が、ほかの学部・学科の学生さんの数学の能力値と同程度になった。そういうふうにとらえていただければというふうに思います。

### 分析結果（物理）：2017年度1年生VS2018年2年生

カリキュラム	能力値 $\theta$ 1年	能力値 $\theta$ 2年	能力値 $\theta$ 差	$p$	$d$
全	0.13	0.19	0.07	0.03	0.07
A	0.16	0.33	0.17	0.03	0.20
B	0.34	0.20	-0.14	0.23	-0.14
C	-0.52	-0.55	-0.03	0.78	-0.04
D	0.12	0.28	0.16	0.07	0.19
E	0.04	0.11	0.08	0.32	0.09
F	0.43	0.39	-0.04	0.68	-0.05
G	0.53	0.68	0.15	0.11	0.19
H	0.43	-0.06	-0.49	0.05	-0.56
I	0.14	0.34	0.19	0.31	0.21
J	-0.46	-0.37	0.09	0.23	0.12

✓ 16年度1年生12月時点の  $\theta$ 平均 = 0  $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		

このほか、簡単にご紹介していきますが、物理の結果ですね。やっぱり白いところが目立ちます。カリキュラムH、これは物理を専門としないカリキュラムなんですけど、これについては有意ですね。能力値が低下したということが分かりました。

### 分析結果（数的文章理解）：2017年度1年生VS2018年2年生

カリキュラム	能力値 $\theta$ 1年	能力値 $\theta$ 2年	能力値 $\theta$ 差	$p$	$d$
全	-0.04	0.00	0.04	0.18	0.04
A	0.15	0.07	-0.08	0.33	-0.09
B	0.45	0.28	-0.18	0.08	-0.22
C	-0.30	-0.06	0.24	0.05	0.40
D	-0.17	0.02	0.19	0.08	0.22
E	-0.08	0.00	0.08	0.40	0.09
F	0.00	0.12	0.12	0.22	0.13
G	-0.04	0.03	0.07	0.53	0.07
H	0.35	0.40	0.05	0.87	0.06
I	-0.22	-0.17	0.06	0.73	0.06
J	-0.03	-0.12	-0.08	0.37	-0.09
K	-0.35	-0.07	0.28	0.05	0.32
L	-0.41	-0.48	-0.07	0.71	-0.08
M	-0.10	0.00	0.10	0.32	0.11
N	-0.07	-0.11	-0.04	0.76	-0.05
O	-0.27	-0.30	-0.04	0.77	-0.05

✓ 16年度1年生12月時点の  $\theta$ 平均 = 0  $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		

数的文章理解はこのような形になっていまして、一部ですね、効果量小で伸びたカリキュラムもあったんですが、全体的に見てあまり変化がなかったという結果になっています。

### 分析結果（化学）：2017年度1年生VS2018年1年生

カリキュラム	能力値 $\theta_{17年}$	能力値 $\theta_{18年}$	能力値 $\theta_{差}$	$p$	$d$
全	-0.20	-0.13	0.07	0.03	0.09
A	-0.12	0.05	0.17	0.04	0.21
B	0.40	0.44	0.05	0.65	0.06
C	-0.45	-0.56	-0.11	0.43	-0.15
D	-0.26	-0.12	0.13	0.15	0.18
E	-0.08	-0.02	0.06	0.51	0.08
F	-0.42	-0.23	0.19	0.02	0.26
G	-0.20	-0.41	-0.21	0.02	-0.29
H	-0.33	-0.37	-0.04	0.83	-0.06
I	-0.43	-0.51	-0.09	0.55	-0.12
J	-0.42	-0.22	0.20	0.01	0.28

✓ 16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		



なんですけれども、これはカリキュラム A・F・G・J について、有意に上がったたり下がったりしていません。

続きまして、分析の2種類目なんですけれども、先ほどは、昨年度の1年生と今年度の2年生を比較したんですけれども、今度は1年生で定点観測したわけです。そういった分析です。昨年度の1年生と今年度の1年生で差があるのかどうかということを分析しました。これについてはまだあまり解釈ができておりませんので、簡単にいまご説明するにとどめたいと思います。まず化学から

### 分析結果（生物）：2017年度1年生VS2018年1年生

カリキュラム	能力値 $\theta_{17年}$	能力値 $\theta_{18年}$	能力値 $\theta_{差}$	$p$	$d$
全	-0.17	-0.17	0.00	0.99	0.00
A	-0.22	-0.12	0.10	0.34	0.09
B	0.51	0.29	-0.22	0.11	-0.21
C	0.07	0.32	0.25	0.11	0.30
D	-0.54	-0.50	0.04	0.72	0.04
E	-0.06	-0.15	-0.09	0.44	-0.09
F	-0.62	-0.53	0.08	0.40	0.10
G	-0.53	-0.62	-0.08	0.43	-0.10
H	-0.37	-0.67	-0.30	0.12	-0.41
I	-0.71	-0.52	0.19	0.27	0.23
J	0.40	0.43	0.04	0.74	0.04

✓ 16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		



生物については、今年度の1年生と昨年度の1年生ではあまり変わりはありませんでした。

### 分析結果（数学）：2017年度1年生VS2018年1年生

カリキュラム	能力値 $\theta_{17年}$	能力値 $\theta_{18年}$	能力値 $\theta_{差}$	$p$	$d$
全	-0.12	-0.08	0.04	0.30	0.04
A	0.03	0.13	0.10	0.19	0.13
B	0.43	0.64	0.21	0.02	0.31
C	-0.95	-0.59	0.35	0.01	0.52
D	-0.20	-0.23	-0.03	0.77	-0.04
E	-0.34	-0.31	0.03	0.79	0.03
F	0.02	0.02	0.00	0.98	0.00
G	-0.03	-0.15	-0.12	0.23	-0.14
H	-0.03	-0.20	-0.18	0.40	-0.22
I	0.04	-0.08	-0.12	0.40	-0.17
J	-0.46	-0.40	0.06	0.46	0.08

✓ 16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

✓  $p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		



数学についてはこのカリキュラム B・C で、効果量小中でそれぞれ上がっている。

### 分析結果（物理）：2017年度1年生VS2018年1年生

カリキュラム	能力値 $\theta_{17年}$	能力値 $\theta_{18年}$	能力値 $\theta_{差}$	$p$	$d$
全	0.09	0.23	0.15	0.00	0.16
A	0.16	0.26	0.10	0.26	0.11
B	0.25	0.82	0.57	0.00	0.61
C	-0.55	-0.33	0.22	0.11	0.30
D	0.05	0.12	0.07	0.51	0.08
E	-0.06	0.05	0.11	0.30	0.12
F	0.41	0.45	0.05	0.61	0.06
G	0.48	0.55	0.07	0.51	0.08
H	0.38	0.52	0.14	0.55	0.15
I	0.15	0.25	0.10	0.62	0.10
J	-0.48	-0.31	0.17	0.04	0.23

$\checkmark$  16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

$\checkmark p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		



物理についても、一部のカリキュラムで上がっているところがあると。

### 分析結果（数的文章理解）：2017年度1年生VS2018年1年生

カリキュラム	能力値 $\theta_{17年}$	能力値 $\theta_{18年}$	能力値 $\theta_{差}$	$p$	$d$
全	-0.08	0.01	0.10	0.00	0.10
A	0.14	0.41	0.27	0.00	0.32
B	0.32	0.79	0.47	0.00	0.56
C	-0.23	-0.02	0.21	0.10	0.31
D	-0.28	-0.60	-0.32	0.01	-0.32
E	-0.16	-0.42	-0.26	0.03	-0.26
F	-0.04	-0.01	0.03	0.81	0.03
G	-0.03	-0.23	-0.20	0.05	-0.23
H	0.28	0.07	-0.21	0.34	-0.25
I	-0.16	-0.02	0.14	0.37	0.18
J	-0.03	0.15	0.18	0.07	0.20
K	-0.31	0.14	0.46	0.00	0.52
L	-0.50	-0.23	0.28	0.15	0.31
M	-0.18	0.05	0.23	0.01	0.28
N	-0.15	-0.01	0.14	0.30	0.16
O	-0.31	-0.14	0.17	0.16	0.21

$\checkmark$  16年度1年生  
12月時点の  
 $\theta$ 平均 = 0  
 $\theta$ 標準偏差 = 1

$\checkmark p < 0.05$  で有意

効果量 $d$	増	減
小		
中		
大		



数的文章理解については、上がっているカリキュラムもあれば、下がっているカリキュラムもあると。こういった状況なんですけれども、この1年生の比較ですね。これについては、少しまだ我々の検討が必要というふうに考えておまして、やはりその、去年の1年生と今年の1年生を比較するということでは、やはり受験に何か要因があるというふうに考えられますので、

例えば昨年の1年生と今年の1年生で、センター試験とか二次試験の成績が上がったか下がったかとか、そういったことが学部学科の先生から伺う必要がありますし、もしくは、何か受験で変えたところがある可能性がありますので、そういったことがなかったかということをご確認させていただいて、データの解釈を今後進めていきたいというふうに思っております。

## まとめ

- 第1回(2017年), および第2回(2018年)の学問基盤力テストの結果を比較分析した
  - 17年度1年生と18年度2年生の比較
  - 17年度1年生と18年度1年生の比較
- 分析の結果, 一部のカリキュラムの学生については平均能力値が有意に向上, あるいは低下したことが明らかになった



ということで、そろそろお時間になりましたので、まとめさせていただきますというふうに思います。私の報告では、昨年度行いました第1回のテスト、そして今年度の4月に行いました第2回のテストも、比較分析の結果をご紹介いたしました。メインの分析になったのは、昨年度に入学された学生さんが、この1年間、昨年度1年間ですね、本学の教育を受けて、その能力値を伸ばしたのかどう

うかと、そういったことに関する分析でした。その分析の結果、一部のカリキュラムの学生さんについては、平均能力値が有意に向上あるいは低下したといったことが明らかになったわけなんですけれども、最初の基盤力テストの目的に立ち返りますと、その目的というのは、本学の学士課程教育の各段階において、その学生の学修達成度を可視化するということでしたので、今後この分析によって、その目的等も一部は達成できてきているんじゃないかというふうに思っております。

## 今後の調査・分析

1. 効果的な学習パターンの探索
  - 学生の能力値の変化と履修歴および学習習慣の関係の分析
2. 基盤力テストの妥当性の検証
  - 各科目の能力値と、概念指標のスコアや特定の授業群のGPAとの相関の分析
  - 項目難易度・能力値の推定法、出題アルゴリズム、能力値のレベル分け基準についての継続的な検討



一応最後の最後なんですけれども、今後の調査・分析としましては、このような形ですね。効果的な解析を行ったりとか、あとはその基盤力テストの妥当性の検証ですね。まだ昨年度と今年度の2回行っただけなんですけれども、この基盤力テストのシステムの改善については、今後も続けていく必要があるというふうに考えております。はい。それでは、お時間となりました

たので、私からの報告は以上とさせていただきます。どうぞご清聴ありがとうございました。

○司会

安田先生、ありがとうございました。具体的な分析結果に加え、テストの考え方など、やや難しいお話になってしまったところもごさいますが、もし、事実確認などごさいましたら、質問お受けできると思います。いかがでしょう。はい、ではお願いいたします。



○質問者 1

非常に興味深い取り組みありがとうございます。やはり気になるのは、テストの時間であるとか、問題数なんですけれども、確かに問題数少ないと、割と簡単にできて、ちょっとした時間で出来るというふうになるんですけれども、例えば教科の中でも単元ぐら이었다ら、こういうような項目で難易度を変えて、非常に有効な方法だと思うんですけれども、ひとつ、例えば物理だったら、いろんな統計がございまして、こういうふうなから5問で問題統制ができていいのかというのを伺いたいですけど。

○安田准教授

非常によい質問ありがとうございます。現時点では、そうですね。科目として、例えば物理の場合ですと、物理というふうにご提示しているんですけれども、実際のところは、力学の分野のみ出題しております。今後そのほかの分野ですと、電磁気とか、あとは波動であったりとか、ほかの分野についてもテストを開発していきたいというふうには考えております。やはりその先生ご指摘のように、何て言うんですかね、5問では物理全体の概念理解については測定出来ないだろうと考えております。今後、第2のセット・第3のセットというふうにご増やしていったら、物理全体の概念理解が測れるような、テストを開発していきたいと考えております。