

平成 30 年 6 月 6 日現在

機関番号：11501

研究種目：挑戦的萌芽研究

研究期間：2016～2017

課題番号：16K12748

研究課題名(和文)大学自然科学系授業のアクティブ・ラーニングに対応した科学概念調査票の開発研究

研究課題名(英文)Scientific Concept Inventory in University Science Curricula

研究代表者

千代 勝実 (Senyo, Katsumi)

山形大学・学士課程基盤教育機構・教授

研究者番号：80324391

交付決定額(研究期間全体)：(直接経費) 2,400,000円

研究成果の概要(和文)：アクティブ・ラーニング型授業等、新しいタイプの実践的な授業の推進が求められているが、これをベンチマーキングし評価・改善を行っていくための適切かつ簡便な科学概念・推論力調査票をスマートフォンのアプリケーションとして開発し実施した。試験は項目反応理論を用い、学生の能力に適切な問題を出題する仕組みを構築した。山形大学の1年生入学当初、2年生4月当初に理系の全学生各1200人に対して数学・物理・化学・生物学の達成度調査を行った。簡易的にセンター試験の受験教科と比較し、蓋然性を確認した。

研究成果の概要(英文)：We have produced standardized tests to investigate and assess students' acquisition of scientific concepts in first year and undergraduate level for benchmarking new teaching paradigms such as the active learning and project-based learning. Tests on smartphone application consist of items implemented according to item response theory, adaptive difficulty for each student. The tests among Mathematics, Physics, Chemistry and Biology were carried out for all freshmen and sophomore, 1200 students each in the beginning of April. Brief consistency check was performed in comparison to National Center Test for University Admissions and standardization.

研究分野：高等教育

キーワード：達成度テスト スマートフォン 項目反応理論 質保証

1. 研究開始当初の背景

この研究課題では、大学の初年次教育など基礎化学・生物学教育における科学的思考力・推論力評価の方法論を確立し、科学概念・推論力調査票を作成し、日本の大学生によって標準化し試行した。

アクティブ・ラーニング型授業等、新しいタイプの実践的な授業の推進が求められているが、これをベンチマーキングし評価・改善を行っていくための適切かつ簡便な指標がこれまでなかった。教育改善のPDCAサイクルを回すために、FCI(力学概念調査)やローソンテストのように、30分程度のペーパーテストでプレテスト・ポストテストを行い、教育効果のゲインを測定できる科学概念・推論力調査票を作成し、日本の大学生で標準化を行うことを当初の目的とした。

研究の背景として、現在の大学教育改革・教学マネジメントにおいて、PDCAサイクルに基づく計画・実施・評価・改善の流れを形成することが求められてきた。計画や実施については各大学で精力的に取り組まれているものの、評価を元にした改善は、教員・学生・外部有識者による主観的な評価が中心であり、エビデンスとして弱いため学内のコンセンサスを形成しがたく改善と新たな企画につなげていくことが難しい。このため教育改善が「PD」の繰り返しに終始している。

一方、教育の中身に目を転じると、教科書を理解し演習問題を解けるだけでなく、概念を理解し実践的に応用できる能力の涵養が求められている。大学における物理学・化学・生物学といった自然科学教育では他分野に比較するとペーパーテストによって定量的に教育効果を測定しやすいという幻想が



あった。しかし教科書を理解し演習問題を解けるようになることが、必ずしも研究や現実世界の問題発見・解決につながらないという反省から、世界的にアクティブ・ラーニングの様々な実践が推進されている(E.F.レディッシュ「科学をどう教えるか」など)。ただ、従来のテストでは演習問題を解く能力しか測れず、面接やレポート・行動観察では客観的なエビデンスを得ることが難しい。

教育改善においてPDCAサイクルが回らない、概念理解や実践力を客観的に測定できないということは同じ問題の2つの側面であり、アクティブ・ラーニング型授業のような

現代的な大学教育の効果測定において適切な評価方法がないということに帰着される。逆に言うと、教育改善には適切な教育効果の評価指標と方法が必須である。

大学の理系基礎科目では、国際的なデファクトスタンダードとして、高等物理教育では物理教育研究に基づいたFCI(力学概念調査票)等、中等教育の理科的推論能力においてはローソンテスト等が存在するが、日本の大学生向けの化学・生物学に関する概念理解および推論能力においては個別の実践例はあるものの自由かつ共通に利用できるものとしては存在しない。

以上のような課題を解決することを企図した。

2. 研究の目的

この研究課題では、以上のような現状を鑑み、FCIおよびローソンテストを参考に、大学初等化学(物理化学)および初等生物学(分子・細胞生物学)における教育効果の指標となる概念および推論能力調査票の作成および標準化までを行うこととした。また企画当初はペーパーテストで実施することとしたが、実際に研究が進むにつれ、スマートフォンベースのアプリケーションに項目反応理論を用いたCBTで実施することとなった。

(1) 本研究は、物理学のFCI(力学概念調査票)のように、30分程度のペーパーテストで、大学理系基礎教育での初等化学(物理化学)、初等生物学(分子・細胞生物学)におけるアクティブ・ラーニングを含めた様々な教育実践の効果測定を行える標準的な概念・推論力調査票を作成し教育効果の測定指標とするものである。これにより大規模調査による従来型授業や様々なアクティブ・ラーニング型授業の教育効果の測定が行える。

アクティブ・ラーニングでの教育効果の測定は、研究の事例毎に制作された評価用の質問紙、もしくはループリックや行動観察、学生への自己評価アンケートなどで評価されることが多い。個別制作された小テストの場合、学習者が理解しているかどうか適切に弁別できる問題を経験的に制作するため、標準化されることはなく他の実践例と直接比較することができない。またループリック等の他者・自己観察によるものは定性的もしくは主観的な評価であるため、測定の安定性という意味で科学的なエビデンスとするには難しい。特に学生の自己評価アンケートは自己洞察やメタ認知能力が高まることによりポストテストで下がることもある。

本研究では、個別の教育実践の評価・調査票ではなく、教育改善需要が大きく見込める大学初等化学と初等生物学に絞って、ベンチマークとしての科学的概念・推論力調査票を制作する。FCI(やローソンテスト)同様、

プレテスト、教育実践、ポストテストの順に実施し、点数の増分（実際は規格化ゲインと呼ばれる量）を教育効果とするものである。これによりある単元を教育する際に、従来型の講義と様々なアクティブ・ラーニング型講義での教育効果についてメタアナリシスの評価が可能になる。

例えば発生に関する従来型授業もしくはアクティブ・ラーニング型授業を行う前後に「モンゴロイド男女のカップルの受精卵を代理母としてコーカソイド女性の子宮に移植した場合生まれてくる子供の特徴は？」といった現実に関わる問題を含む調査票を解かせて同じ指標の元で教育効果を測定し授業内容の改善につなげていく。

(2) 本研究では、概念調査票を作成する際に化学や生物学教育におけるベストプラクティスとその具体的な設問を収集しメタアナリシス的な検討を加えるため、必然的に「よい教育のアウトカム」とは何なのかを明示するとともに、教育手法横断的な評価を理工系教育に導入する。

本研究の初等化学（物理化学）、初等生物学（分子・細胞生物学）での標準的な概念調査票を制作するという目的上、これまでの従来型講義とアクティブ・ラーニング型授業で想定されてきた具体的な教育目的や評価基準における学習者の「よい理解」「よい態度」と定義されているものを収集し、それらを過不足なく概念・推論力調査票上で表現することになる。これは、その設問自体が良いのか悪いのかという問いを通じて、本来その科目で教育すべきことは何なのかという本質的な質問を、同じ概念・推論力調査票という平面的な俎上に載せることになる。また、さまざまな教育実践の先行研究から設問意図を収集することにより、集合知としてのあるべき教育の方向性というのが浮き彫りになってくる。

物理学では FCI 等の設問表現なども含めた妥当性の検証を通して、教育すべきこと、すべきでないことを明確にしてきた。またその結果学生が持っていた物理は暗記科目という考え、教員が持っていた古典的な演習問題をたくさん課せば概念理解につながるという考えを正し、アクティブ・ラーニング型授業の優位性を明らかにしてきた。（新田 日本科学教育学会 2015）

この影響が他の理工系専門科目へ波及したのはいままでのことであるが、日本における化学や生物学教育においても、本研究で制作する科学概念・推論力調査票により「教育すべきこと」「教育すべきでないこと」に関して再定義と明示化を行い、アクティブ・ラーニング型授業や従来型授業における定量的なエビデンスを得ることにより、教育改善に資することができる。

さらにこの概念・推論力調査票への批判や

検証による改良、またより優れた手法による評価指標の開発に関する興味が学会や教育界で喚起されると考える。

3. 研究の方法

当研究課題は化学・生物学における「科学的概念・推論力の定義」「科学的概念・推論力調査票の作成」「大学生による指標の標準化と最適化」を行い、最終的に科学的概念・推論力調査票作成とベンチマーク化による統計解析手法の確立を目指す。まず、基本的な調査と測定すべき学力の定義づけ、調査票のための問題収集を中心的行い、少数の被験者のもとで調査票の完成を目指す。その後、調査票の完成と評価、これによって得られたデータから統計解析に必要なサンプル数や精度についてシミュレーションを実施し推定する。被験者 1000 人規模で複数大学での調査を導入し解析を行うとともに課題の洗い出しと、実際の授業実践を効果測定することによる精度の向上を行う。

(1) 研究の背景

物理学における力学概念調査 FCI の作成（ヘステネス 1992）、およびその妥当性の検証という似た先行事例があるため、本課題における化学・生物学におけるテストの作成とその検証については、それをトレースして開発研究する。我々はすでに他プロジェクトで FCI を用いて物理学における概念・推論力に関わるアクティブ・ラーニング型の教育研究を実施し、また設問の妥当性調査を行っている。この取り組みを化学・生物学の調査票作成の手本とする。また FCI の 5 者択一方式（ただし正しい組み合わせを選ぶといった設問もある）を基準に、評価の定量化と単純化という観点を可能な限り導入する。

(2) 具体的な研究・手法

調査票開発の手順は 大学初年次の学生の持つ素朴概念や、化学・生物学を学び現実に適用する上での科学的推論力を、教員・学生へのインタビューや先行研究の質問票から収集し、それを元にテストのひな形を作成し 1 教科 100 人程度の被験者に受験させ難易度と点数分布を確認しつつ、1 教科 10 人程度の被験者に対しインタビューなどの方法で解答する際の思考過程を明確化し、表現方法や選択肢の順番などを含めた整理をし、調査因子の重複や欠損を整理し設問数を極小化し、100 点中平均点が 60 点程度となり標準偏差が 15 点程度になるような規格化をおこなう。最終的にリハーサルをおこない本実施となる。これらの開発研究においては化学、生物学、全体のチェックと統計処理をそれぞれの専門分野について個別に作成する。

科学的思考力・推論力の定義・収集においては、先行研究の質問票に加え理学部・工学

部・医学部・農学部
の各教員からも聞き
取りをおこない、初
年次教育および専門
教育における知見を
集約する。

初期調査として理
系学部だけでなく文
系学部の学生にも広
くテストを受験させ、
その解答状況から設
問表現についてもな
るべく専門用語を排
除し日常的な科学的
思考の態度が発現し
やすい質問を開発す
る。同時に解答に至
った理由も記述式で
収集する。

少数の被験者で、
解答過程を観察しイ
ンタビューによりな
ぜその解答に至った
のか、またクロスフ
ィード（間違っただ
けで正しい解答に至
る、もしくは正しい理解で間違っただけで至る）となる要因を検出する。

「必ず」「全て」という単語を忌避するよ
うな文脈依存性や正解の選択肢に特定の数字
が続く（例えば正答の選択肢がずっと2など）
と忌避するような解答テクニク的な偏りを抽出
し、逆にそれらの設問への利用可能性を含めて
検討する。文化的・教育制度上の要因も排除する。

調査されるべき科学的思考力・推論力につ
いての整理をおこない設問数の最小化をおこ
なう。

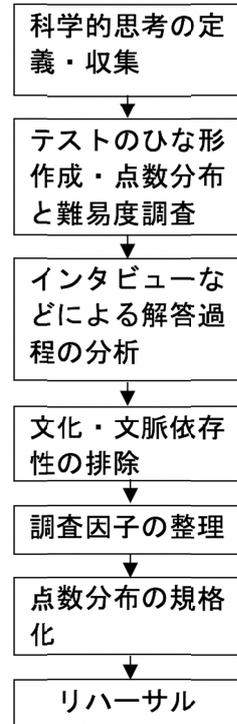
調査感度が最大になるように点数分布の
規格化を実施する。これについては での結果
を検討し100名程度の被験者に実際に受験し
てもらふことにより難易度とその得点分布を
調整する。

以上の方針でテストの作成は、内外の状況を
調査しつつ基礎化学（物理化学）と基礎生物
学（分子・細胞生物学）の開発研究をおこな
う。

調査票の作成でポイントとなるのは、科学
的概念・推論力の定着度を正しく測定するた
め学生の態度や行動を答案に顕在化できる
か、という点である。物理学においては設問
の構成についておおよそのコンセンサスが
あるが、化学・生物学については当研究課題
が萌芽的な研究であるため確固としたもの
はない。設問を構成する要素の抽出が具体化
できない場合は早期に設問検討会を開催し、
早期に適切な設問群の構成をおこなう。

科学的概念・推論力調査票の開発研究は、

テスト作成の過程



化学・生物学について、点数の規格化および1000人程度の被験者によるリハーサルを実施する見込みである。また、この際、山形大学だけではなく、名古屋大学および中部大学の協力を得て、さまざまな学習背景を持つ学生にリハーサルを行う。これは物理学などのアクティブ・ラーニング型授業研究です。名古屋大学など東海圏の大学の教員および高等教育研究センターと協力関係にあるため、教養・共通教育や理系基礎科目における科学的概念・推論調査票や調査手法に関するノウハウが共有されているからである。リハーサルでは、化学および生物の化学概念・推論調査票を利用し各大学の協力を得てデータの取得を行う。研究代表者・分担者（千代：とりまとめ、飯島：化学、渡辺：生物学、安田：データ分析）および協力大学の担当者が講義内で実際に1年次入学当初、もしくは個別の授業開始時のプレテスト・講義終了後のポストテストをおこない、大学・学部・学科ごとの平均および分布、ポストテスト後に変化の度合いを教育効果として定義し、科学的概念・推論力を統計処理して学生個人の達成度と授業のアウトカムを導出し、指標の標準化を行う。さらにこれらの結果から分析結果の解釈を行う。

また、教育改善のPDCAサイクルを考えた

調査票設問の例

（化学、概念理解型設問）

- 次のうち、化学変化が起きたものを全て選べ。
(1) 水が冷えて氷ができた。
(2) 鉄くぎが錆びた。
(3) 電池を使って豆電球を点灯させた。
(4) 海水から水が蒸発し塩が析出した。
(5) 大豆が発酵して納豆ができた。

（生物学、科学的推論型設問）

- 黒髪のモンゴロイド夫婦間の子どもを金髪のコーカソイド女性が代理母として出産した場合、正しいものを選べ。
(1) 子どもも、将来生まれる孫も黒髪になる。
(2) 子どもが黒髪、金髪になる割合は3:1である。
(3) 子どもは黒髪であるが、子どもが黒髪の黄色人種と結婚しても金髪の子どもが生まれる可能性がある。
(4) 代理母の持つ遺伝子型により異なるため、分らない

きに、複数科目が含まれるカリキュラム全体の教育効果を評価する手法を検討する。リハーサルにより得られたデータの解析結果や分布を使用し、過去の講義の受講学生の学部・学科・性別の平均的な人数分布情報からモンテカルロシミュレーションを反復して

おこない、個々の授業実践に対して達成度や教育効果を有意に示すためにはどの程度のサンプル数が必要になるかを探る。

4. 研究成果

調査票の開発を28年度に完成した。手順は4～10月に大学初年次の学生の持つ素朴概念や、教科を学び現実に適用する上での科学的思考力を収集し、それを元にテストのひな形を作成し12月に1教科100人程度の被験者に受験させ難易度と点数分布を確認しつつ、1教科10人程度の被験者に対しインタビューなどの方法で解答する際の思考過程を明確化し、表現方法や選択肢の順番などを含めた整理をし、調査因子の重複や欠損を整理し設問数を極小化し、100点中平均点が60点程度となり標準偏差が15点程度になるような規格化をおこなった。最終的に1月にリハーサルをおこない平成29年度に本実施となるよう、準備を行った。これらの開発研究においては物理学、化学、生物学を分担者で分担し、交付申請書に記載の科目に加えて、数学、数的文章理解をそれぞれの専門分野について個別に作成した。また、実際のテストの実施においては、テスト出題に学生保有のスマートフォンを使用し、回収や採点・集計の手間を省くとともに、項目反応理論を用い、学生の学力にあわせて出題する設問の難易度を変化させ、学生の達成度を測定するために必要となる解答数や解答時間を最小化した。これにより1教科平均5分程度の時間で推定できる調査票を作成した。

テストの実施について、平成29年4月5日、11日に、作成した調査票をスマートフォンベースのアプリケーションで実施した。数的文章理解については山形大学の1年次全学生(約1700名)、自然科学系科目では山形大学1年次の対象学生(理系・約1200名)に対して、大学1年次レベルの数学・物理学・化学・生物学の問題で99%以上の学生に受験させることができた。この問題は、特に単に情報や知識の量を問うのではなく、広い意味でのアクティブラーニングを通じて効果的に獲得されると考えられる、科学概念や応用力、環境への相互作用など実践的な能力が測定できているか、を確認する問題である。分析結果では、どの科目も学生集団が平均値に対して概ね ± 2 標準偏差の中に収まっており、また1年終了時の学生に比べて平均値は概ね下回っていた。暫定的な検証であるが、この差が山形大学での1年間の学習達成度の増加分ではないかと考えている。また、入学時の準備度を検証するために、生物において、センター試験で生物を受験した学生とそうでない学生に分類し、調査票の得点分布を比較すると、明確に分布の上位はセンター試験で生物を受験した学生群で占められた。このため、センター試験と我々の作成した概念理解を問う調査票のスコアには相関が存在すると

考えられる。

研究期間外となったが、平成30年4月3～6日において、2年次(29年度1年次で受験した学生層)と平成30年度の1年次の学生に同様のテストを行い、2年次の学生では約85%、1年時の学生では99%以上の受験率で実施した(平成30年4月26日現在)。このデータを基に1年間で学生の達成度がどのように変化したのかを分析するとともに、理系科目やそれ以外の授業・授業外学修などの状況との相関、さらに他大学でのデータ収集と評価を今後継続していく予定である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 0 件)

[学会発表](計 9 件)

安田淳一郎、基盤力テストの現状と今後の展望、APキックオフシンポジウム及び第10回EM-IR勉強会、明治大学、2016年

安田淳一郎、山形大学における基盤力テストの現状と課題、大学教育改革フォーラム in 東海2017、金城学院大学、2017
千代勝実、安田淳一郎、渡辺絵理子、飯島隆広、山形大学における基盤力テスト～全学横断的な達成度テストの概要～、日本物理学会、岩手大学、2017

安田淳一郎、千代勝実、渡辺絵理子、飯島隆広、山形大学における基盤力テスト～第1回実施データの分析～、日本物理学会、岩手大学、2017

千代勝実、基盤力テストの概要と実施状況、APシンポジウム「直接評価の第一歩 基盤力テストの実施と活用に向けた取組」、大正大学、2017

安田淳一郎、分析結果 I: 学問基盤力、APシンポジウム「直接評価の第一歩 基盤力テストの実施と活用に向けた取組」、大正大学、2017

Katsumi Senyo, Shigeru Asano, Koji Fujiwara, Takao Hashizume, Developing a Smartphone-Based Integrated Student Assessment Data Collection System to Enhance Yamagata University's Education, Assessment Institute in Indianapolis(国際学会)、2017

Takao Hashizume, Katsumi Senyo, Shigeru Asano, Koji Fujiwara, Redesigning the FYE Course at Yamagata University: Our Challenging Journey, 37th Annual Conference on The First-Year Experience(国際学会)、2018

千代勝実、物理学講義におけるアクティ

ブラーニングの効用と課題、大学教育改
革フォーラム in 東海 2018、中京大学、
2018

〔その他〕
ホームページ等

<https://ir.yamagata-u.ac.jp>

6．研究組織

(1)研究代表者

千代 勝実 (SENYO, Katsumi)
山形大学・学士課程基盤教育機構・教授
研究者番号：80324391

(2)研究分担者

安田 淳一郎 (YASUDA, Junichiro)
山形大学・学士課程基盤教育機構・准教授
研究者番号：00402446

渡辺 絵理子 (WATANABE, Eriko)
山形大学・学士課程基盤教育機構・准教授
研究者番号：20337405

飯島 隆広 (IJIMA, Takahiro)
山形大学・学士課程基盤教育機構・准教授
研究者番号：20402761