

令和 5 年 6 月 8 日現在

機関番号：11501

研究種目：基盤研究(C)（一般）

研究期間：2019～2022

課題番号：19K03135

研究課題名（和文）部分設問群とマルチステージテストを用いた精確かつ効率的な力学概念評価テストの開発

研究課題名（英文）Developing a Computerized Multistage Test with the Force Concept Inventory

研究代表者

安田 淳一郎（YASUDA, Jun-ichiro）

山形大学・学士課程基盤教育院・准教授

研究者番号：00402446

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究課題の主要な成果は次の2点である。  
力学概念指標（FCI）を用いたコンピュータ適応型テスト（CAT）を開発した。典型的な条件下でシミュレーション分析を実施し、紙面版FCI（全30問）の精度から10%の精度低下を許容すれば、FCI-CATの出題数を15問まで短縮可能なことを明らかにした。

FCI-CATのアルゴリズムを改良し、テストの効率性、および安全性を高めた。同様の分析により、改良版FCI-CATでは、精度低下なしで、出題数を事前・事後合計で29問まで短縮可能なことを明らかにした。また、テストの効率性と安全性のバランスを実現するには、事前・事後の出題数を同程度にすべきことを示した。

研究成果の学術的意義や社会的意義

本研究の成果により、コンピュータ適応型テストを利用することで、力学概念指標（FCI）の実施に要する時間を約30分から約15分まで短縮できることが明らかになった。これにより、これまで調査に費やされていた授業時間を学習時間に充てるのが可能になる。また、これまで時間的な制限からFCI調査を授業中に実施することを躊躇っていた教員にとっては、実施を後押しすることが期待できる。以上のように、本研究で開発されたFCI-CATは学習評価に係わる負担を軽減することにより、評価の普及を促進するものである。

研究成果の概要（英文）：The two main results of this research project are as follows.

(1) We developed computerized adaptive testing (CAT) using the Force Concept Inventory (FCI). Using simulations, we found that for a class size of 40, we may reduce the test length of the FCI-CAT to 15-19 items (the full-length FCI has 30 items) with an accompanying decrease in accuracy and precision of only 5-10%.

(2) We improved the algorithm of the FCI-CAT to increase the efficiency and security of the test. The simulation analysis revealed that we can achieve the accuracy and precision of the full-length FCI with fewer items via the FCI-CAT. Specifically, for a class size of 40, we can reduce the sum of the pre- and post-test lengths of the FCI-CAT to a total of 29 items, thereby reducing the testing time to 55%. We also found that if one desires a balance of test efficiency and security, choosing equal pre- and post-test lengths would be reasonable.

研究分野：物理教育研究

キーワード：物理教育研究 科学教育 教育評価 学習評価 テスト理論 項目反応理論 コンピュータ適応型テスト

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

## 1. 研究開始当初の背景

大学教育において、学修成果の可視化が求められている。中央教育審議会大学分科会の審議まとめ(案)には、「制度改正等の方向」として「学生の学修成果」を「的確に把握・測定し、教育活動の見直し等に適切に活用することが求められる」とあり、学修成果の測定の制度化が進む可能性が高まっている。

物理教育の学修成果を可視化する上で重要な問いの1つが、学習者の物理概念の理解度をいかに正確に測定するかという問いである。つまり、理解度の評価テストの測定誤差をいかに小さくし、信頼性・妥当性をいかに高めるかという問いである。代表的な研究としては、Hestenesらが開発した力学概念指標 (Force Concept Inventory, FCI) の精確さについての研究が挙げられる。FCI は、力学の概念的理解度を測定するための評価テストであり、多肢選択式の設問 30 問で構成される。FCI の妥当性は、文章や図、誤答選択肢など様々な観点から評価されており、研究代表者らも FCI の偽正答(理解していない状態での正答)に関する妥当性について研究を続けてきた。具体的には、学習者の理解状態を判定するため、ある設問で複合的に問われている複数の概念を単一的な概念に分解し、それぞれの単一的な概念について個別に問う設問群(以下、部分設問群)を独自に開発した。

評価テストの精度を高めることは重要であるが、テストの設問数を増やすと、テストの実施にかかる時間や負担が増大するという、精度と効率性のジレンマの問題がある。この問題をいかに解決するか、つまり、学習者の物理概念の理解度をいかに正確かつ効率的に測定するかという問いが本研究課題の核心をなす問いである。

テストの効率性を高めるための手がかりとなるのが、PISA 等で採用されている、コンピュータ適応型テスト (Computerized Adaptive Testing, CAT) である。CAT では、受験者の解答に応じてコンピュータから出題される設問が最適化される。たとえば、第 1 問に正答した場合は第 2 問により難しい設問が出題される。これにより、テストの精度低下を抑制しつつ、出題数を大幅に減らすことができる。ただし、CAT を単純な多肢選択式の設問で構成した場合は、適応型でない通常のテストと同様に偽正答の影響を大きく受ける可能性がある。

## 2. 研究の目的

本研究の目的は、力学概念に焦点を絞り、学習者の力学概念の理解度を正確かつ効率的に測定するための評価テストを開発することである。開発するテストの土台として、FCI を用いる。テストの精度を高めるため、テストに FCI の部分設問群を組み込み、偽正答が生じる確率を低減する。テストの精度を維持しつつ効率性を高めるため、CAT の発展型であるマルチステージテスト (Multistage Testing, MST) を採用する。CAT では 1 問ごとに次問への分岐が生じるが、MST では複数の設問を 1 組にまとめて 1 モジュールとし、1 モジュールごとに次モジュールへの分岐が生じる。本研究での 1 モジュールは、FCI のある設問とその部分設問群で構成する。

## 3. 研究の方法

本研究は以下のような工程で進めた。CAT は MST よりも構造が単純であるため、第 1 段階として FCI を組み込んだ CAT (以下、FCI-CAT) を開発し、システムの理解に必要なデータを収集した。第 2 段階として FCI とその部分設問群を組み込んだ MST (以下、FCI-MST) を開発した。研究期間の途中で COVID-19 の影響により、当初計画していた調査が実施困難になったため、当該期間を利用し、FCI-CAT の最適な出題設問数を分析する研究、および、FCI-CAT のアルゴリズムを改良する研究を実施した。

## 4. 研究成果

(1) 第 1 段階として、FCI-CAT を開発した。開発に際して以下の分析を行った。2015 年から 2018 年に実施した FCI 調査データ (2882 名分) を利用し、CAT の構成に必要な項目反応理論の困難度パラメータを推定した。同推定は項目反応理論の複数のモデルに基づいて試行し、データと各モデルの適合度を調べた。また、FCI が一次元性等の項目反応理論の前提条件を満たすか否かを確認した。

(2) FCI-CAT を実装した。実装においては CAT のオープンソースプラットフォームである Concerto Platform を利用し、受験者がスマートフォンから解答できるようにした。開発完了後、国内 3 大学で試行調査を実施し、結果に基づきシステムを改良した。

(3)シミュレーション分析により、FCI を CAT 化することで、紙面版 FCI 全 30 問と同程度の精度で、出題数をどれほど短縮可能なのかを調べた。手順は次の通りである。モンテカルロ・シミュレーションに基づき、FCI-CAT の擬似解答データを生成した。FCI の実解答データを用いる Post-hoc シミュレーションの結果と、生成した擬似解答データの結果との整合性を確認した。テストの精度を効果量 (Cohens'  $d$ ) の二乗平均平方根誤差で定量化し、テスト精度とテスト出題数の関係性を調べた。分析の結果(図 1), 受講者数が 40 名の場合、5~10%の精度低下を許容すれば、FCI の出題数を 15~19 問(FCI 全 30 問中 50~63%)まで短縮可能なことが明らかになった。本研究成果は、PRPER 誌の注目論文として Editors' Suggestion に選ばれた[1]。

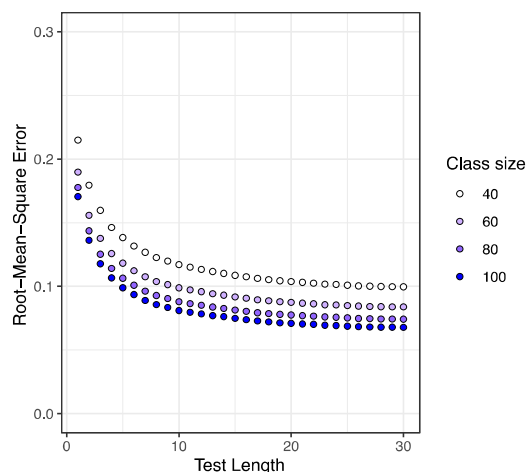


図 1 . FCI-CAT の出題数と精度の関係

(4)FCI-CAT のアルゴリズムを以下の観点から改良した。FCI で分類された物理概念が均等に 出題されるように、出題バランスを制御した CAT アルゴリズムを導入した[2]。事前テストに 関する記憶が事後テストまで保持される可能性を考慮し、設問の露出率を制御する CAT アルゴリズムを導入した。事前テストでの能力推定値を付帯情報として事前分布に組み込み、事後テ ストの項目選択や能力推定に利用するアルゴリズムを導入した。事後テストの出題数を、事前 テストの出題数より短縮する可能性を検討した。

(5)改良した FCI-CAT について、(3)と同様の分析を実施し、出題数をどれほど短縮可能なのかを 調べた[3]。分析の結果(図 2), 受講者数が 40 名の場合、精度低下なしで、出題数を事前・事 後合計で 29 問まで短縮できることが明らかになった。また、テスト効率性とテスト安全性は、 互いに相反し、両者のバランスを考慮するのであれば、事前・事後テストの出題数を同程度にす べきことを示した。

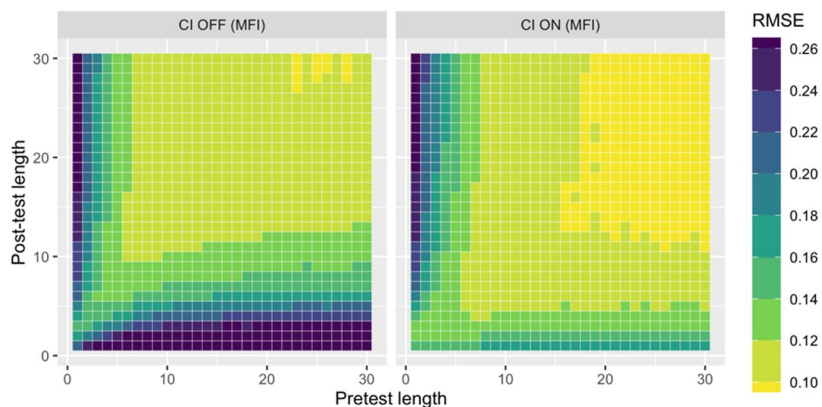


図 2 . 改良した FCI-CAT の出題数と精度の関係。左図が改良前で、右図が改良後。紙面版 FCI の精度 RMSE=0.10 を実現可能な、最小設問数の組合せは、事前：16 問、事後：13 問。

(6)FCI-MST を次の手順で開発した。2022 年に FCI とその部分設問群を含む調査を実施し、そ れらの項目困難度パラメータを推定した。同推定は項目反応理論の複数のモデルに基づいて試 行し、データと各モデルとの適合度を調べた。また、FCI と部分設問群の項目プールが一次元性 等の項目反応理論の前提条件を満たすか否かを確認した。シミュレーションに基づき、FCI- MST における能力値の推定等のアルゴリズムとしてどれが適切かを検討した。

< 引用文献 >

[1] J.Yasuda, N.Mae, M.M.Hull, and M.Taniguchi, "Optimizing the length of computerized adaptive testing for the Force Concept Inventory," Phys. Rev. Phys. Educ. Res. **17**, 010115 (2021)  
 [2] J. Yasuda and M. M. Hull, "Balancing content of computerized adaptive testing for the Force Concept Inventory," 2021 Physics Education Research Conference Proceedings, pp.456-461 (2021)  
 [3] J.Yasuda, M.M.Hull, and N.Mae, "Improving test security and efficiency of computerized adaptive testing for the Force Concept Inventory," Phys. Rev. Phys. Educ. Res. **18**, 010112 (2022)

## 5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 4件/うちオープンアクセス 4件）

1. 著者名 Yasuda Jun-ichiro, Hull Michael M., Mae Naohiro	4. 巻 18
2. 論文標題 Improving test security and efficiency of computerized adaptive testing for the Force Concept Inventory	5. 発行年 2022年
3. 雑誌名 Physical Review Physics Education Research	6. 最初と最後の頁 010112_1-15
掲載論文のDOI（デジタルオブジェクト識別子） 10.1103/PhysRevPhysEducRes.18.010112	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Yasuda Jun-ichiro, Hull Michael M.	4. 巻 2021
2. 論文標題 Balancing content of computerized adaptive testing for the Force Concept Inventory	5. 発行年 2021年
3. 雑誌名 PERC Proceedings	6. 最初と最後の頁 456-461
掲載論文のDOI（デジタルオブジェクト識別子） 10.1119/perc.2021.pr.yasuda	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Yasuda Jun-ichiro, Mae Naohiro, Hull Michael M., Taniguchi Masa-aki	4. 巻 1929
2. 論文標題 Analysis to Develop Computerized Adaptive Testing with the Force Concept Inventory	5. 発行年 2021年
3. 雑誌名 Journal of Physics: Conference Series	6. 最初と最後の頁 012009_1-9
掲載論文のDOI（デジタルオブジェクト識別子） 10.1088/1742-6596/1929/1/012009	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Jun-ichiro Yasuda, Naohiro Mae, Michael M. Hull, Masa-aki Taniguchi	4. 巻 17
2. 論文標題 Optimizing the length of computerized adaptive testing for the Force Concept Inventory	5. 発行年 2021年
3. 雑誌名 Physical Review Physics Education Research	6. 最初と最後の頁 010115_1-15
掲載論文のDOI（デジタルオブジェクト識別子） 10.1103/PhysRevPhysEducRes.17.010115	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計12件（うち招待講演 3件 / うち国際学会 3件）

1. 発表者名 安田 淳一郎, Michael M. Hull, 前 直弘
2. 発表標題 冲積図と部分設問群を用いた力学概念理解の系統的分析
3. 学会等名 日本物理学会2023年春季大会
4. 発表年 2023年

1. 発表者名 YASUDA Jun-ichiro
2. 発表標題 Optimizing the length of computerized adaptive testing for the Force Concept Inventory
3. 学会等名 University of Alaska Fairbanks, Spring 2023 Journal Club (招待講演)
4. 発表年 2023年

1. 発表者名 安田 淳一郎
2. 発表標題 コンピューター適応型テストによる学習成果の可視化：理論と活用
3. 学会等名 九州大学 基幹教育院 粒子系理論物理学研究室セミナー（招待講演）
4. 発表年 2023年

1. 発表者名 安田 淳一郎, Michael M. Hull, 前 直弘
2. 発表標題 力学概念指標を用いたコンピュータ適応型テストの開発 IV
3. 学会等名 日本物理学会2022年秋季大会
4. 発表年 2022年

1. 発表者名 YASUDA Jun-ichiro, HULL M. Michael
2. 発表標題 Balancing content of computerized adaptive testing for the Force Concept Inventory
3. 学会等名 Physics Education Research Conference 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 YASUDA Jun-ichiro
2. 発表標題 Optimizing the length of computerized adaptive testing for the Force Concept Inventory
3. 学会等名 2021 AAPT Virtual Summer Meeting (国際学会)
4. 発表年 2021年

1. 発表者名 YASUDA Jun-ichiro
2. 発表標題 Optimizing the length of computerized adaptive testing for the Force Concept Inventory
3. 学会等名 Austrian Educational Competence Center, University of Vienna, 520001 SE Design Based Research-new developments (Online) (招待講演)
4. 発表年 2021年

1. 発表者名 安田 淳一郎, 前 直弘, Michael M. Hull, 谷口 正明
2. 発表標題 力学概念指標を用いたコンピュータ適応型テストの開発III
3. 学会等名 日本物理学会第76回年次大会
4. 発表年 2021年

1. 発表者名 安田 淳一郎, 前 直弘, Michael M. Hull, 谷口 正明
2. 発表標題 力学概念指標を用いたコンピュータ適応型テストの開発II
3. 学会等名 日本物理学会2020年秋季大会
4. 発表年 2020年～2021年

1. 発表者名 安田 淳一郎, 前 直弘, Michael M. Hull, 谷口 正明
2. 発表標題 力学概念指標を用いたコンピュータ適応型テストの開発
3. 学会等名 日本物理学会 2019年 秋季大会
4. 発表年 2019年

1. 発表者名 安田 淳一郎, 前 直弘, Michael M. Hull, 谷口 正明
2. 発表標題 力学概念指標を用いたコンピュータ適応型テストの開発
3. 学会等名 第36回物理教育研究大会
4. 発表年 2019年

1. 発表者名 YASUDA Jun-ichiro, MAE Naohiro, HULL M. Michael, TANIGUCHI Masa-aki
2. 発表標題 Analysis to Develop Computerized Adaptive Testing with the Force Concept Inventory
3. 学会等名 GIREP-ICPE-EPEC-MPTL 2019 (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	前 直弘 (MAE Naohiro)  (10796098)	大阪大学・核物理研究センター・協同研究員  (14401)	
研究分担者	谷口 正明 (TANIGUCHI Masa-aki)  (90554113)	名城大学・その他部局等・教授  (33919)	

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	ハル マイケル (HULL M. Michael)	アラスカ大学フェアバンクス校・Assistant professor	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計1件

国際研究集会 FCI-CAT Winter Meeting	開催年 2019年～2019年
----------------------------------	--------------------

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関		
米国	アラスカ大学フェアバンクス校		
オーストリア	ウィーン大学		