

令和元年6月7日現在

機関番号：11501

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K00948

研究課題名(和文)力学概念指標の系統的誤差研究に基づく妥当性評価のための新しい聞き取り調査法の開発

研究課題名(英文) Developing an interview method based on the systematic error of the Force Concept Inventory

研究代表者

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

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：学習者の力学概念を測定するための指標としてForce Concept Inventory(以下、FCI)が国際的に広く使われている。本研究では、FCIの系統的誤差を分析するため、日米の大学、延べ16大学において質問紙調査を実施し、2226名の学生から回答を得た。主たる成果に関する分析として、FCI問5、6、7、16の偽正答(誤った理由で正答すること)に起因する系統的誤差の和を求め、FCI30問の系統的誤差について最小限の値を推定した。その結果、スコア中位の集団については、偽正答に起因する系統的誤差の大きさが少なくとも生スコアの10%に達することが明らかになった。

研究成果の学術的意義や社会的意義

Force Concept Inventory (FCI) の系統的誤差を求めようとする本研究の試みは、国外を含めて前例がなく、世界的に独創的なアプローチである。本研究の結果、FCIの1つの系統的誤差の値が部分的に明らかになったことから、従来の統計的誤差の分析と合わせ、学習成果の分析をより精密に行うための一歩を踏み出すことができた。これにより、物理教育研究の手法が一般的な自然科学研究の手法にまた一歩近づくことになり、ひいては科学教育分野の信用を高めることが期待できる。本研究の成果は、精密な測定が可能なテストの開発において役に立つ可能性がある。

研究成果の概要(英文)：The Force Concept Inventory (FCI) is a widely used assessment test to probe student conceptual understanding of Newtonian mechanics. In order to analyze the systematic error of the FCI, we surveyed students at 16 (in total) universities in Japan and the United States. The total number of survey responses was 2226. As the main analysis, we calculated the sum of the systematic errors due to the false positives from Q.5, Q.6, Q.7, and Q.16 of the FCI. Then, we estimated a minimum value for the systematic error due to false positives on the entire FCI. We find that the sum of the false positives from Q.5, Q.6, Q.7, and Q.16 is about 10% of the FCI score of a mid-level student.

研究分野：科学教育

キーワード：物理教育研究 科学教育 教育評価 学習評価 概念指標

様式 C-19、F-19-1、Z-19、CK-19 (共通)

1. 研究開始当初の背景

物理学の学習成果を測定するためのアセスメント・テストの中で、国際的に最も広く普及しているテストの一つが力学概念指標 (Force Concept Inventory, FCI) である [1]。FCI は学習者の Newton 的力学概念の理解度を測定するためのテストであり、五肢択一式の設問 30 問で構成される。これまでに FCI は、Hake [2] によるアクティブラーニング型授業の効果検証など、様々な教授法の効果検証において重要な役割を果たしてきた。

FCI の妥当性の評価、すなわち、FCI が学習者の Newton 的力学概念の理解度を本当に測定できているか否かの評価は、様々な観点から行われている。たとえば、Hestenes らは学生と教員への聞き取り調査によって、解答者が FCI の設問文と設問の図を正しく理解していることを確かめている [1, 3]。また Stewart らは、ある設問で問われている物理概念は変えず、設問の文脈だけを変えた設問を作成し、変更前後で解答者の正答率を調べることで、FCI の正答率が設問の文脈にそれほど依存しないことを確かめている [4]。

FCI の妥当性は様々な観点から確認されているが、FCI で生じる偽正答については議論の余地がある。偽正答とは、解答者がある設問で問われている物理概念を理解していないにもかかわらず、正答することである。たとえば、FCI の設問は五肢択一式であるため、解答者が無作為に選択肢を選んだ場合に 20% の確率で偽正答が生じる。FCI 開発者の Hestenes らは、「FCI には素朴概念をもつ解答者を惹きつける誤答選択肢が備わっているため、偽正答は軽減されている」[3] としつつも、学生への聞き取り調査において偽正答が「かなりよく生じた」[1] としている。さらには、複数の研究者は FCI 問 16 で偽正答が特に生じやすいことを指摘している [5-8]。具体的には、問 16 は Newton の第 3 法則を適用して解くべき設問でありながら、Newton の第 1 法則を適用して正答する解答者が一定数いることが指摘されている。

偽正答は FCI の生得点を真得点から押し上げるため、FCI の系統的誤差になり得る。FCI の偽正答に起因する系統的誤差については、Hake [2] によって簡単に考察されているものの、詳細については十分に調べられてこなかった。特に、問 16 で生じる特異な偽正答が近年になって指摘されたにもかかわらず、その系統的誤差については、ほとんど調べられてこなかった。

以上のような問題意識から、研究代表者らは FCI の偽正答に起因する系統的誤差について研究を進めてきていたが、本研究課題の開始当初には次のような問題があった。(1) 前研究では、FCI 数問の分析結果から FCI 全体の系統的誤差を概算していたが、概算の際の仮定の妥当性が確かでなかったため、分析方法を改善する必要があった。(2) 系統的誤差を求める際に、偽正答の影響のみ考慮しており、偽誤答の影響は含めていなかった。偽正答には生得点を押し上げる効果があるが、偽誤答には押し下げる効果があるため、系統的誤差がいくぶん縮小する可能性があった。(3) Hake [2] が指摘するように、系統的誤差がたとえ大きかったとしても、平均規格化ゲインや Cohen's d 等の効果量を計算する際に、事前テストの平均点と事後テストの平均点の差の計算において、系統的誤差が相殺される可能性があった。

2. 研究の目的

本研究の目的は、前節で示した問題に対応させると、次の 3 点にまとめられる。(1) FCI 数問の分析結果から FCI 全体の系統的誤差を概算する方法を改め、FCI の偽正答に起因する系統的誤差を再評価する。(2) 偽誤答に起因する系統的誤差を求め、偽正答に起因する系統的誤差と合算することにより、FCI の系統的誤差について、より精確な値を求める。(3) FCI の偽正答および偽誤答に起因する系統的誤差が平均規格化ゲインや Cohen's d 等の効果量に与える影響を評価する。

3. 研究の方法

本研究において重要な工程となる、学習者の偽正答・偽誤答の判定においては、安田・谷口 [7] が考案した部分設問群の方法を用いた。部分設問群とは、FCI のある設問で問われている複数の力学概念を個々の力学概念に分解し、各々の力学概念の理解を個別に問えるようにした設問群のことである。部分設問群を用いた判定においては、FCI のある設問を正答した場合に、その設問の部分設問群を 1 つでも誤答した場合は、偽正答と判定する。一方、FCI のある設問を誤答した場合に、その設問の部分設問群をすべて正答した場合は、偽誤答と判定する。

(1) FCI の偽正答に起因する系統的誤差を評価するための調査では、FCI 問 5, 6, 7, 16 各設問の部分設問群および FCI 全 30 問によって調査票を構成した。問 6, 7, 16 は安田・植松・新田 [6] の聞き取り調査で偽正答が見出されていた設問である。これらの設問の比較対象として、同調査で偽正答が見出されていなかった問 5 も調査票に含めた。分析では、2015 年 4 月、2017 年 4 月に国内の大学 4 大学において実施した調査のデータを用いた。本調査の解答者数は 1145 人であり、有効解答者数は 1110 人であった。

前研究では、FCI 数問の分析結果から FCI 全体の系統的誤差を概算する方法に問題があった。そこで、本研究では個々の設問の系統的誤差を個別に求め、それらの和を FCI 全体の系統的誤差の最小限とする分析方法を考案した。個々の設問の系統的誤差を個別に求める方法を考案したことの副次的な成果として、個々の設問の系統的誤差を比較することが可能になった。

(2) FCI の偽誤答に起因する系統的誤差を評価するための調査では、FCI 問 13, 18, 26, 27 各設問の部分設問群および FCI 全 30 問によって調査票を構成した。問 18, 26, 27 は安田・植松・新田 [6] の聞き取り調査で偽誤答が見出されていた設問である。これらの設問の比較対象として、同調査で偽正答が見出されていなかった問 13 も調査票に含めた。調査は 2016 年 4 月、2018 年 4 月に国内

の大学 5 大学において実施した。本調査の解答者数は 1446 人であり、有効回答者数は 1266 人であった。

偽誤答に起因する系統的誤差の分析においても、個々の設問の系統的誤差を個別に求め、それらの和を FCI 全体の系統的誤差の最小限とする分析方法を採用した。

(3) FCI の偽正答・偽誤答に起因する系統的誤差が平均規格化ゲインおよび Cohen's d 等の効果量に与える影響を評価するため、シミュレーションを用いて、事前テストおよび事後テスト各々について 10,000 人分の解答データを生成した。そのデータを用いて、(系統的誤差を考慮しない)生の効果量と、(系統的誤差を考慮した)真の効果量を求め、値の比較を行った。

4. 研究成果

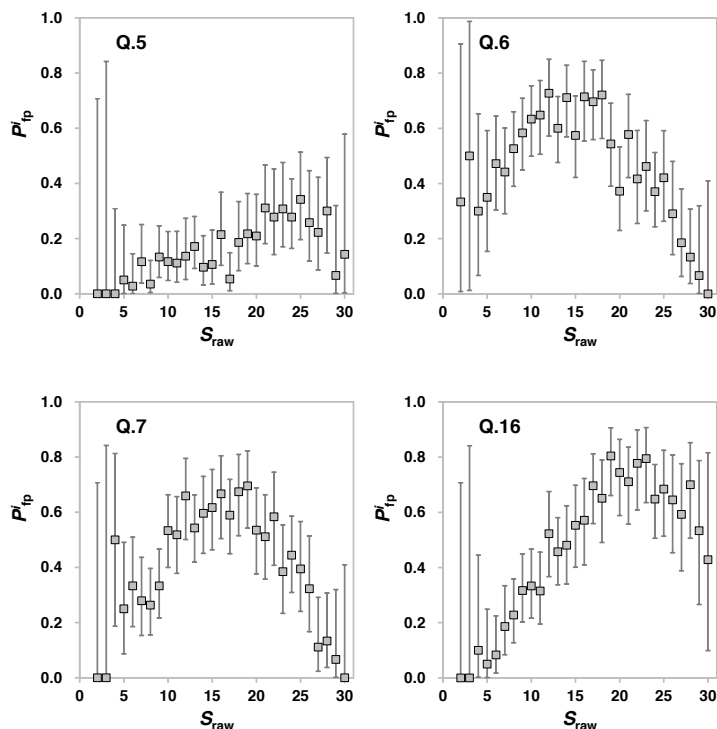


図 1 生得点 (S_{raw}) と偽正答に起因する系統的誤差 (P_{fp}^i) の関係 (問 5, 6, 7, 16)

(1-①) FCI 問 5, 6, 7, 16 について、偽正答に起因する系統的誤差を設問ごとに分析した [11]。図 1 の横軸 (S_{raw}) は FCI30 問の生得点、縦軸 (P_{fp}^i) は各設問の系統的誤差 (生正答率と真正答率の差) を表す。図 1 より、偽正答に起因する系統的誤差は低・高得点層と比べ、中得点層において特に大きいことが明らかになった。また、安田・植松・新田 [6] の聞き取り調査で偽正答が見出された設問 (問 6, 7, 16) の系統的誤差は、偽正答が見出されていない設問 (問 5) の系統的誤差よりも大きいことが明らかになり、聞き取り調査と質問紙調査の結果の一貫性が確認された。

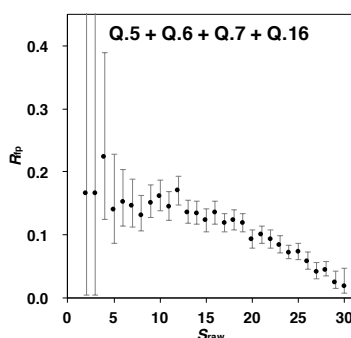


図 2 生得点 (S_{raw}) と、系統的誤差の和と生得点の比 (R_{fp}) の関係 (問 5, 6, 7, 16)

(1-②) 問 5, 6, 7, 16 の系統的誤差の和を求め、それと生得点との比を分析した [11]。図 2 の横軸 (S_{raw}) は FCI30 問の生得点、縦軸 (R_{fp}) は系統的誤差の和と生得点の比 ($\sum_{i=5,6,7,16} P_{fp}^i / S_{raw}$) を表す。図 2 より、系統的誤差は中得点層において大きい値を取り、生得点の 10% に達することが明らかになった。分析した 4 問のみの系統的誤差で 10% に達したため、FCI 全 30 問の系統的誤差は 10% よりも大きくなると考えられる。また図 2 より、系統的誤差は中得点層において大きい値を取るが、高得点になるに従い、系統的誤差は小さくなることも明らかになった。満点に近い解答者については、偽正答をほとんどしていないことも明らかになった。

(2-①) 部分設問群を作成した 8 問 (問 5, 6, 7, 16 および, 問 13, 18, 26, 27) について, 偽正答に起因する系統的誤差と偽誤答に起因する系統的誤差の値を比較したところ, 偽誤答に起因する系統的誤差は偽正答に起因する系統的誤差の約 2 割にとどまることが明らかになった。この結果は, 偽正答が生じる確率について「かなりよく生じた」[1] としていた一方, 偽誤答が生じる確率について「10%より小さい」[3]としていた Hestenes らの分析結果と矛盾しなかった。

(2-②) FCI 問 5, 6, 7, 16 および問 13, 18, 26, 27 について, 偽正答・偽誤答に起因する系統的誤差の和を求め, それと生得点との比を分析した。分析の結果, 系統的誤差は中得点層において大きい値を取り, 生得点の 10%に達することが明らかになった。また, 系統的誤差は中得点層において大きい値を取るが, 高得点になるに従い, その値は小さくなることも明らかになった。つまり, 偽誤答に起因する系統的誤差を合算しても, 図 2 で示した結果からほとんど変わらないことが明らかになった。

(3) シミュレーションで生成した解答データ (事前テスト, 事後テスト各 10,000 人分) を用いて, (系統的誤差を考慮しない) 生の効果量と, (系統的誤差を考慮した) 真の効果量を求め, 値の比較を行った。事前テストの生の平均点が 12.4 のとき, 真の平均点は 10.8 となり, 系統的誤差の影響によって 13.1%の差が生じた。また, 事後テストの生の平均点が 15.7 のとき, 真の平均点は 13.9 となり, 系統的誤差の影響によって 11.3%の差が生じた。これらの値に基づき, 平均規格化ゲインを計算したところ, 生の値は 0.186, 真の値は 0.167 となり, 9.87%の差が生じた。一方, Cohen's d を計算したところ, 生の値は 0.578, 真の値は 0.580 となり, 0.34%の差しか生じなかった。このことから, 効果量の計算においては, Cohen's d の方が偽正答および偽誤答に起因する系統的誤差の影響を受けにくいいため, 教授法の教育効果の分析においては Cohen's d の使用が推奨されることが示唆された。

<引用文献>

- [1] D. Hestenes, M. Wells, and G. Swackhamer, *Phys. Teach.* **30**, 141 (1992).
- [2] R. Hake, *Am. J. Phys.* **66**, 64 (1998).
- [3] D. Hestenes and I. Halloun, *Phys. Teach.* **33**, 502 (1995).
- [4] J. Stewart, H. Griffin, and G. Stewart, *Phys. Rev. ST Phys. Educ. Res.* **3**, 010102 (2007).
- [5] R. K. Thornton, D. Kuhl, K. Cummings, and J. Marx, *Phys. Rev. ST Phys. Educ. Res.* **5**, 010105 (2009).
- [6] J. Yasuda, H. Uematsu, and H. Nitta, *J. Phys. Educ. Soc. Jpn.* **59**, 90 (2011).
- [7] J. Yasuda and M. Taniguchi, *Phys. Rev. ST Phys. Educ. Res.* **9**, 10113 (2013).
- [8] M. Taniguchi and J. Yasuda, *J. Phys. Educ. Soc. Jpn.* **62**, 226 (2014).
- [9] J. Yasuda, N. Mae, M. M. Hull, and M. Taniguchi, *Phys. Rev. Phys. Educ. Res.* **14**, 010112, 2018.

5. 主な発表論文等

[雑誌論文] (計 3 件)

- ① J. Yasuda, N. Mae, M. M. Hull, and M. Taniguchi, “Analyzing the Measurement Error from False Positives in the Force Concept Inventory,” *Proceedings of GIREP-MPTL 2018*. (査読有, 受理済)
- ② J. Yasuda, N. Mae, M. M. Hull, and M. Taniguchi “Analyzing False Positives of Four Questions in the Force Concept Inventory,” *Phys. Rev. Phys. Educ. Res.* **14**, 010112, American Physical Society, 2018. (査読有)
DOI: <https://doi.org/10.1103/PhysRevPhysEduRes.14.010112>
- ③ M. M. Hull, J. Yasuda, M. Taniguchi, and N. Mae “Towards Quantification of the FCI's Validity: the Effect of False Positives,” *2017 Physics Education Research Conference Proceedings*, pp. 180-183, 2018. (査読有)
DOI: 10.1119/perc.2017.pr.040

[学会発表] (計 6 件)

- ① 前直弘「FCI の系統的誤差が教育効果の分析に与える影響」日本物理学会第 74 回年次大会, 九州大学, 2019 年 3 月.
- ② J. Yasuda, “Analyzing the Measurement Error from False Positives in the Force Concept Inventory,” GIREP-MPTL 2018, San Sebastian, Spain, Jul. 2018.
- ③ 安田淳一郎「力学概念指標の偽正答に起因する系統的誤差の Conservative な推定」日本物理学会第 73 回年次大会, 東京理科大学, 2018 年 3 月.
- ④ M. M. Hull, “Towards Quantification of the FCI's Validity: the Effect of False Positives,” *Physic Education Research Conference 2017*, Cincinnati, Ohio, Jul. 2017.
- ⑤ 安田淳一郎「力学概念指標の偽誤答に起因する系統的誤差分析」日本物理学会 2016 年秋季大会, 金沢大学, 2016 年 9 月.
- ⑥ J. Yasuda, “Estimating Systematic Error from False Negatives of the Force Concept Inventory,” World Conference on Physics Education 2016, São Paulo, Brazil, Jul. 2016.

6. 研究組織

(1)研究分担者

研究分担者氏名:谷口 正明

ローマ字氏名:(TANIGUCHI, masa-aki)

所属研究機関名:名城大学

部局名:教職センター

職名:准教授

研究者番号(8桁):90554113

研究分担者氏名:前 直弘

ローマ字氏名:(MAE, Naohiro)

所属研究機関名:関西大学

部局名:システム理工学部

職名:特別任用教授

研究者番号(8桁):10796098

(2)研究協力者

研究協力者氏名:Michael M. Hull

所属研究機関名:ウィーン大学

※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、研究者個人に帰属されます。