

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 13 日現在

機関番号：11501

研究種目：若手研究(B)

研究期間：2014～2015

課題番号：26750068

研究課題名(和文) 部分設問群を用いた系統誤差評価に基づく国際的通用性のある教育効果分析法の構築

研究課題名(英文) Developing a method of educational evaluation focused on a systematic error using subquestions

研究代表者

安田 淳一郎 (YASUDA, JUNICHIRO)

山形大学・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) 1,000,000円

研究成果の概要(和文)：本研究では、FCIの妥当性を系統的誤差として定量化し、その振る舞いを分析した。我々が着目した系統的誤差は、偽正答に由来するものである。偽正答の判定にあたっては、部分設問群を用いた。部分設問群を作成した設問には、以前の聞き取り調査で偽正答が見出された問6, 7, 16に加えて、調査対象としなかった設問についての概算のために問5を含めた。分析の結果、FCIの得点およびその平均規格化ゲインについての系統的誤差は真値と同程度の値に達することを明らかにした。また、問6, 7, 16についての真正答率とFCI30問の生正答率の関係において閾値が存在することを明らかにした。

研究成果の概要(英文)：In this study, we quantified the validity of FCI as a systematic error and analyzed its behavior. We focused on the false positives and used subquestions to judge whether a student understands the concept tested in an FCI question. The questions for which we created the subquestions are limited to the three invalid-like questions that, according to prior research, elicit false positives from students, and one of the remaining 27 valid-like questions. We used the results of the valid-like question and its subquestions to estimate the systematic error arising from the remaining 26 valid-like questions. Our results indicate that true score can be lower than half of raw score and true average normalized gain can be as low as half of raw average normalized gain, depending on the fitting function for the estimation. We also found that there are thresholds in the relationship between true positive ratio and raw score for the invalid-like questions, but not for the valid-like question.

研究分野：物理教育研究

キーワード：学習評価 概念調査 教育評価 力学概念 妥当性評価 系統誤差

1. 研究開始当初の背景

物理教育の効果測定を目的として、様々な指標の研究開発が国内外で活発に行われている。中でも、最も盛んに研究されている指標の一つが力学概念指標 (Force Concept Inventory, FCI) である¹⁾。FCIは、学習者の力学概念を測定するための指標であり、五肢択一の設定問30問で構成されている(図1)。これまでFCIは31カ国語に翻訳されており、世界的に最も広く普及している指標の1つと言える。国内では、東京理科大・東京学芸大グループ等がFCIを和訳するなど、FCIを用いた物理教育研究が広まりつつある²⁾。

FCI等の指標を用いる場合、その妥当性を調べる必要がある。一般に、妥当性とは調査の目的とその手段の整合性のことであり、FCIの妥当性評価では、「FCIの設定問等が学習者の力学概念を正確に測定できるよう構成されているか」を調べる。これまでにFCIの妥当性は、文章や図の妥当性³⁾、選択肢の妥当性⁴⁾、設問と物理概念との対応関係の妥当性⁴⁾など、様々な観点から評価されている。

代表研究者らも、FCIの妥当性評価に関して、顕著な研究成果をあげてきた。最初の研究⁵⁾では、事前にFCIに回答した大学生16名を対象に「どの選択肢を選んだのか」「なぜその選択肢を選んだのか」等を尋ねる半構成型聞き取り調査を行った。その結果、一部の学生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(偽正答)していることを見出した。また、一部の学生はある設問で問われている物理概念を理解しているにもかかわらず誤答(偽誤答)しているも見出した(表1)。特に、ある設問においては、特定の誤った理由で正答する回答者の割合が非常に大きいという、重大な欠陥を

見した。ただし、調査の被験者数が十分でなかったため、見出したFCIの欠陥の確からしさについて統計学的な分析を行うには至らなかった。

次の研究では、統計学的に十分な信頼度確からしさでの分析ができるよう、質問紙調査を実施した⁶⁾。分析対象とするFCIの設定問は、聞き取り調査の結果に基づき、偽回答が現れる可能性の高い2問に絞り込んだ。そして、それらの設問について、偽回答を判定するために必要な思考過程の単位まで設問を細分化した設問群(以下、部分設問群)を独自に開発した。調査は国内1大学で実施し、111件の回答を得た。分析の結果、「調査対象とした2問は妥当でないと95%の確からしさで言える」ことが明らかになった。この結論は、その後の補完的な調査においても変わらなかった。

しかし、同研究のアプローチにも限界があった。FCIのある設問が妥当か否か、を評価する方法では、妥当でない一部の設問がFCI全30問の総得点(全正答率)に与える影響を定量的に評価できなかった。また、仮に同研究の結果を踏まえ、FCIの妥当でない一部の設問を修正した場合、修正前のFCIの調査結果と修正後のFCIの調査結果を比較することが困難になるという問題があった。

そこで代表研究者らは、一部の設問の欠陥を定量化し、FCI全30問の総得点(全正答率)の系統的誤差として扱えないかと考えた。このアプローチにより、現行のFCIの設定問を修正する必要がなくなり、これまでに蓄積されたFCIの調査結果を有効に活用できるようになる。

このアイデアに基づき分析を行った結果、聞き取り調査で偽正答が見出されたFCIの一部の設問の系統的誤差は、通常の設定問の系統的誤差の4~7倍程度に及ぶこと、および、その系統的誤差の値は統計的誤差の値と同程度の大きさであることが明らかになった。

2. 研究の目的

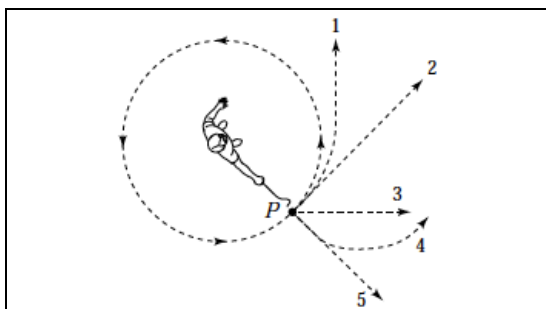
本研究の目的は、代表研究者らがこれまでに明らかにしたFCIの系統的誤差の現実性を高め、FCIの不確かさを取り込んだ教育効果分析法を構築することである。さらには、系統的誤差の特性について国際的な共通性や相違点を明らかにし、構築した教育効果分析法の国際的通用性を高めることを目指す。

本研究の副次的な成果として、妥当性の高い指標を開発するための手がかりを得ることが期待できる。また、本研究では学習者の認知過程を精密に分析することから、学習者の認知過程に関する新たな法則を発見できる可能性もある。

3. 研究の方法

(1) 調査票の開発

部分設問群の例として、図1で示したFCIの設定問の部分設問群を図2に示す。部分設問群は、FCI1問につき3問程度で構成した。調



【問題】図のように、鉄球にひもをつけて、水平面上で円を描くように回転させる。図のP地点で、ひもが切れたとき、その後の球の進路に最も近いものを選びなさい。
図1. FCIの設定問例(概略のみ)

	理解している	理解していない
正答	真正答	偽正答
誤答	偽誤答	真誤答

表1. 各行は、ある設問に正答したか否かを表す。各列は、その設問で問われている物理概念を理解しているか否かを表す。

査票の開発にあたっては、回答者が最後まで集中して取り組めるよう、総設問数を40問程度に絞り込んだ。

図2の部分設問群では、FCIの設問例(図1)を解くために必要なすべての物理概念を問うている。そのため、被験者が部分設問群をすべて正答したか否かによって、被験者が設問で問われている物理概念を理解しているか否かを判定することができる。つまり、偽回答の判定は、表1に対応づけると、表2のようにして行うことができる。

調査対象としたFCIの設問は、FCI問5, 6, 7, 16の4問である。これらの設問のうち問6, 7, 16については、以前の聞き取り調査で偽正答が現れた設問であり、系統的誤差が特に大きいことが予想される設問である。問5は聞き取り調査で偽正答が現れなかった設問であるが、比較分析のために、また調査対象としなかった残り26問についての系統的誤差を概算するために調査対象とした。

さらに、生正答率ごとの系統的誤差の振る舞いを分析するために、部分設問群に加えて、FCI30問を調査票に含めた。これで総設問数は40問となった。

(2) 調査の実施

調査は2015年4月に、日本の国立大1大学、私立大3大学の物理学に関連するコースの授業中に実施した。有効回答数は502件であった。ほとんどの回答者は1年生であり、所属は理系の学部生を中心に多岐に渡った。

(3) 調査結果の分析

① 部分設問群の文脈や選択肢等に微小な変

【問題】ひもが切れた直後の、次の変数の向きを、1~8から選びなさい。「水平方向に働く力はゼロ」と考える場合には、9と記入しなさい。

1. ボールにはたらく力
2. ボールの加速度
3. ボールの速度

図2. 設問例(図1)の部分設問群(概略のみ)

	部分設問群を全て正答	部分設問群を1つ以上誤答
正答	真正答	偽正答
誤答	偽誤答	真誤答

表2. 各行は、ある設問に正答したか否かを表す。各列は、その設問の部分設問群を全て正答したか否かを表す。

化(ゆらぎ)を与え、その系統的誤差への影響を調べることにより、部分設問群の妥当性を評価した。

② FCI30問の単純な正答率(生正答率)ごとに回答者を層別に分け、各層について真正答率を求めた。最初に、各設問(問5, 6, 7, 16)についてFCI30問の生正答率と各設問の真正答率の関係を分析した。次に、FCI30問についての生正答率とFCI30問の真正答率の関係を分析した。FCI30問の真正答率は、FCI30問の生正答率と問5の真正答率の関係について回帰分析した結果を用いて概算した。FCI30問の生正答率とFCI30問の真正答率の関係は、一般的な測定装置についての特性曲線に対応する。FCIの特性曲線と、FCIが偽正答の観点から完全に妥当である場合の理想的な直線(生正答率=真正答率)の差を求め、その差をFCIの系統的誤差と同定した。

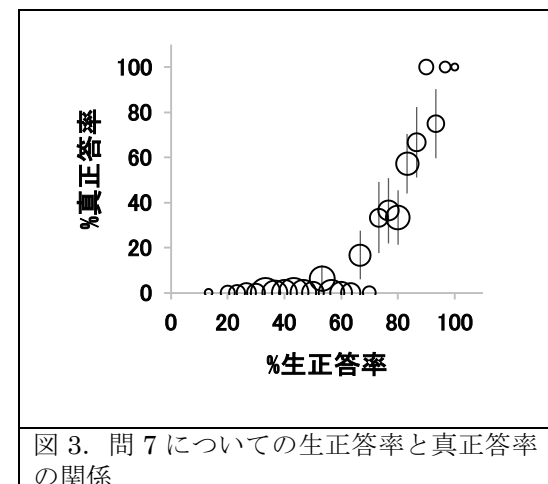
③ FCIの系統的誤差が教育方法の効果を表す平均規格化ゲイン⁷⁾など、正答率についての複合的な変数に与える影響を調べ、分析②と同様に平均規格化ゲインの系統的誤差を求めた。

4. 研究成果

【研究の主な成果】

(1) 前研究から本研究に移行する過程において、部分設問群の選択肢に与えたわずかな変化が偽正答率に与える影響を分析し、その結果がこれまでの分析結果と矛盾がないことを確認した⁸⁾。これにより、部分設問群のわずかな違いによる系統的誤差への影響は十分に小さいことの傍証を得た。

(2) FCI問6, 7, 16についての真正答率とFCI30問の生正答率の関係において閾値が存在することを明らかにした。この結果について、問7を例に説明する。図3は、FCI30問について、ある生正答率を得た回答者ごとに算出した、問7の真正答率の振る舞いを表す。ここで、真正答率は(設問の真正答者数÷設問の生正答者数×100)で定義される。各バブルのサイズが大きいほど、その生正答率を獲得した回答者数が多いことを表す。またバブルを貫く線分は真正答率についての標準誤差を表す。グラフより、生正答率が約60%より



小さい、ほとんどの回答者層の真正答率はゼロである一方、生正答率が約 60%より大きい回答者層については、生正答率が増加するにつれて真正答率も増加し、真正答率=100%に達していることがわかる。生正答率 =60%に閾値があるというこの結果は、Hestenes らの分析結果と一致している⁹⁾。

(3) FCI30 問についての系統的誤差が真値と同等の値に達する場合があることを明らかにした。図 4 の横軸はある回答者の FCI30 問についての単純な正答率を、縦軸は部分設問群を含めて算出した真正答率を表す。図中の破線は、生正答率と真正答率が一致し、FCI がこの分析の観点から完全に妥当である場合を表す。図中の実線は、回帰分析から求めた曲線(特性曲線に対応)であり、調査で得られた生正答率と真正答率の関係を表す。FCI の系統的誤差とは、その破線と実線の間のギャップに相当する。数値を見ると、生正答率が 50%のとき、真正答率はおよそ 25%であり、系統的誤差の値が真値に相当する場合があることがわかる。

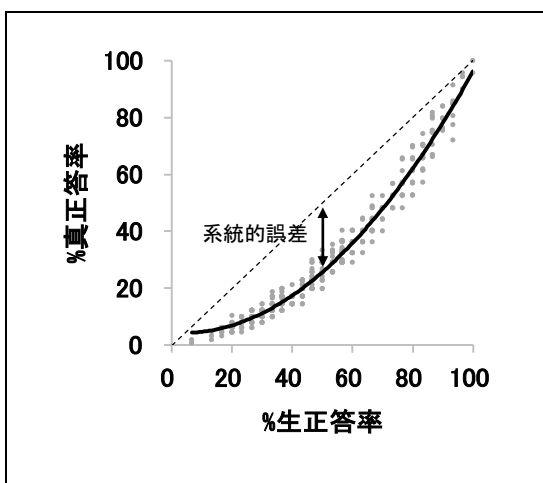


図 4. FCI30 問の生正答率と真正答率の関係

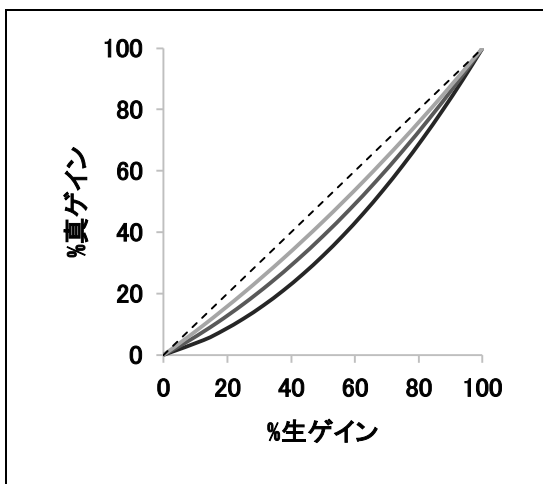


図 5. 生ゲインと真ゲインの関係についてのシミュレーション結果

(4) 平均規格化ゲインについての系統的誤差が真値と同等の値に達する場合があることを明らかにした。図 5 の横軸はある回答者の単純な平均規格化ゲインを、縦軸は部分設問群を含めて算出した真の平均規格化ゲインを表す。図中の破線は、生ゲインと真ゲインが一致し、FCI がこの分析の観点から完全に妥当である場合を表す。図中の実線は、事前生スコア=20% (薄灰色)、40% (灰色)、60% (濃灰色) それぞれの場合について生ゲインおよび真ゲインの値をシミュレートした結果である。平均規格化ゲインの系統的誤差は、破線と各実線の間のギャップに相当する。数値を見ると、事前生スコア=60%の場合に生ゲインが 50%のとき、真ゲインはおよそ 25%であり、平均規格化ゲインについての系統的誤差の値が真値に相当する場合があることがわかる。

【得られた成果の国内外における位置づけとインパクト】

本研究の成果は、3回の国内学会および2回の国外学会において、全て口頭発表によって報告を行った。また、本研究の成果をまとめた論文を2015年12月に、Physical Review Special Topics Physics Education Research誌に投稿し、現在査読を受けている最中である。これらのことより、本研究の成果は国内外の研究コミュニティにおいて一定のインパクトを与えたものと考えられる。

国内で行った調査と同様の調査は、Wayne State CollegeのMichael Hull氏の協力を得て、2015年8月に米国で実施した。ただ、その被験者数は100名弱と、当初予定していた被験者数を下回ったため、統計学的な分析を行うには至っていない。Hull氏からは継続的な協力を得ることについて内諾を得ているため、今後も調査を継続し、十分な量のデータを収集することにより、本研究で得られた系統的誤差についての国際通用性を高めていく予定である。

【今後の展望】

これまでの系統的誤差の分析は、次の点で未だ不十分である。〔1〕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い。偽正答は真正答率を押し下げる効果を持つが、偽誤答は押し上げる効果をもつため、上記の系統的誤差がいくぶん縮小する可能性がある。〔2〕上記の分析では、FCI数問についての調査結果からFCI全体の系統的誤差を単純に概算しているに過ぎない。より精確な系統的誤差の値を求めるためには、その概算法を洗練する必要がある。〔3〕これまでの主な調査対象は日本の学生であるため、系統的誤差の国際的一般性を確かめるためには、米国など国外の学生を対象にさらに調査を行う必要がある。

さらに、系統的誤差の研究を進める中で、次のような新たな問いも浮かんでいる。5つ

の選択肢からなる FCI の設問では、回答者が無作為に回答しても 20%の確率で偽正答が生じる。そのような無作為選択による偽正答と、特定の誤った理由で正答することで生じる深刻な偽正答では、系統的誤差への寄与の違いが3倍にも達することが研究代表者らの分析で明らかになっている。これは、その深刻な偽正答を取り除けば、系統的誤差を小さくできることを意味する。深刻な偽正答を取り除くための方法としては聞き取り調査があり、FCIを開発した Hestenes とも聞き取り調査を行っていた。しかし実際には、深刻な偽正答を取り除けなかったということは、その聞き取り調査の方法に何らかの不備があることを意味している。そこで今後の研究では、系統的誤差に関する研究成果に基づき、様々な物理概念指標の妥当性評価において用いることのできる新しい聞き取り調査の方法の開発を予定している。

引用文献

- 1) D.Hestenes, M.Wells and G.Swachkamer: Phys.Teach., 30-3(1992)141-158.
- 2) 新田英雄, 塚本浩司: 大学の物理教育, 17-1(2011)16-19.
- 3) N.Rebello and D.Zollman: Am.J.Phys. 72-1(2004)116-125.
- 4) T.F.Scott, D.Schumayer and A.R.Gray: Phys.Rev.ST Phys.Educ.Res.8(2012)020105.
- 5) 安田淳一郎, 植松晴子, 新田英雄: 物理教育, 59-2(2011)90-95.
- 6) J.Yasuda and M.Taniguchi: Phys. Rev. ST Phys. Educ. Res., 9, 010113, (2013).
- 7) R.Hake: Am.J.Phys.66(1998)64.
- 8) 谷口正明, 安田淳一郎: 物理教育, 62, pp.226-231, (2014).
- 9) D. Hestenes and M. Wells, Phys. Teach. 30, 159 (1992).

5. 主な発表論文等

〔雑誌論文〕(計1件)

- ① 谷口正明, 安田淳一郎「部分設問群を用いた和訳版 FCI の定量的な妥当性評価」物理教育, 第 62 巻, 第 4 号, pp.226-231, 日本物理教育学会, 2014 年(査読有).
<http://ci.nii.ac.jp/naid/110009896270>

〔学会発表〕(計5件)

- ① 安田淳一郎, 谷口正明, 前直弘, M.M.Hull 「力学概念指標の偽正答に着目した系統的誤差分析」日本物理学会第 71 年会次大会, 東北学院大学, 2016 年 3 月.

- ② J.Yasuda, M.Taniguchi and N.Mae, “Characteristics of a Systematic Error of the Force Concept Inventory,” International Conference on Physics Education 2015, Beijing Friendship Hotel, Beijing, China, 2015 (査読有).

- ③ 安田淳一郎「力学概念指標の妥当性評価研究」日本物理学会第 70 回年次大会(若手奨励賞受賞記念講演), 早稲田大学, 2015 年 3 月.

- ④ 安田淳一郎, 谷口正明「Subquestions を用いた FCI の系統誤差分析」日本物理学会 2014 年秋季大会, 中部大学, 2014 年 9 月.

- ⑤ J.Yasuda and M.Taniguchi, “Analyzing A Systematic Error of the Force concept inventory with Sub-Questions,” International Conference on Physics Education 2014, “Dr. César Carman” Hotel, Córdoba, Argentina, 2014 (査読有).

〔その他〕(計1件)

- ① 日本物理学会「第 9 回(2015 年)領域 13 若手奨励賞」2015 年 3 月.

6. 研究組織

(1) 研究代表者

安田 淳一郎 (YASUDA, Jun-ichiro)
山形大学・基盤教育院・准教授
研究者番号: 00402446

(2) 研究協力者

谷口 正明 (TANIGUCHI, Masa-aki)
名城大学・教職センター・准教授

前 直弘 (Mae, Naohiro)

関西大学・理工学教育開発センター・特任教授

Mike Hull

Wayne State College, Department of Physical Sciences and Mathematics
Assistant Professor.