

令和 5 年 6 月 8 日現在

機関番号：34504

研究種目：基盤研究(C)（一般）

研究期間：2020～2022

課題番号：20K03081

研究課題名（和文）到達目標型大学英語教育のデザインと評価

研究課題名（英文）An Application of DBR to the Development of a Framework for University English Language Education

研究代表者

住 政二郎（Sumi, Seijiro）

関西学院大学・生命環境学部・教授

研究者番号：60441341

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）：本研究は、デザイン研究のフレームワークに基づき、課題分析と教材・テスト開発を行い、自動問題生成システムに実装し、授業実践に活用することに加え、結果を項目反応理論によって分析しスコアを標準化することで到達目標型大学英語教育の実現を目指すことを目標とした。本研究では、スコアを標準化するためにイロレーティングの教理モデルを活用した。イロレーティングは、チェスプレイヤーの勝敗結果から各プレイヤーのランキングを標準得点で示すために開発された。本研究の結果、少人数クラスを対象に、テスト結果から受験者能力値と困難度パラメータをリアルタイムで把握するシステム開発に成功し、授業実践への応用を実現した。

研究成果の学術的意義や社会的意義

本研究の主要な学術的意義は、イロレーティングを応用し、テスト結果から受験者能力値と困難度パラメータを即座に計算し、リアルタイムでその結果を可視化することに成功したことである。加えて、自動問題生成システムを使い、デザイン研究のフレームワークに従い授業モデルを提案したことである。社会的意義は、開発に成功した自動問題生成システムを公開し、学内の必修英語の授業で実際に活用できるようにしたことである。また、新型コロナウイルス感染拡大を受け、オンラインでもテストを実施できるように改良を加え、テストシステムとしての汎用性を高め、学内以外の教育現場でも活用できるようにしたことである。

研究成果の概要（英文）：The purposes of this study are to develop (a) test items and teaching materials based on the framework for the design-based research, (c) implement them in the Automated Test Maker, which make it possible to generate test sets along with the lessons, and (d) design a comprehensive English lesson. To standardize the test scores, the Elo rating system was used, which was originally developed to rank players at the interval scale in a fair manner. The study made it possible to estimate and monitor students' ability and item difficulty parameters on a real-time basis and apply them to design English lessons.

研究分野：教育工学

キーワード：英語教育 テスト イロレーティング 学習支援システム 授業研究

1. 研究開始当初の背景

研究開始当初、研究と教育の有機的連携を促進し教育実践の改善に取り組む「デザイン研究」(Design-based Research) が注目を集めていた。デザイン研究の歴史は古く、1990年代の初期には Brown や Collins らが「デザイン思考」の重要性について指摘している。IDEO が 1998 年に実施したショッピングカートプロジェクトは、人間の行為を考慮した道具のデザイン・設計が、現場の課題解決にいかに重要かを示した。その後、2003 年には Design-Based Research Collective が当時の研究・実践を総括し、Educational Researcher は 2003 年に、Educational Technology は 2005 年にデザイン研究の特集号を組んだ。van den Akker (2006) は、デザイン研究と教育におけるテクノロジー利用との親和性を早くから指摘し、教育学の研究に多大な影響を与えた。その後、デザイン研究は、Design Research in Education や Educational Design Research など、強調する点の違いにより、多様なアプローチへと発展したが、そのプロセスには、「課題の分析・焦点化」、「開発・実践」、「分析・考察・省察」、「理論的枠組みの提示」の要素をいずれも共通項として持つ (Bakker, 2018)。日本教育工学会 2019 年秋季全国大会では、国内外で活躍する第一線の研究者らを集めてのキーノート・シンポジウムが開催され、今後のさらなる展開が期待されていた。しかし、外国語教育学分野での応用例は極めて少なく、研究成果の蓄積は乏しかった。

外国語能力を測定するテストの歴史も古い (Goundar, 2020; Himelfarb, 2019; Vinovskis, 2019)。特に英語のテストに関しては、グローバル化と項目応答理論の深化に伴い広範なテスト市場を形成するようになった。テスト結果の「標準化」(standardization) と「比較可能性」(comparability) という項目応答理論の特性は、標準テストを世界に広げる強力な要素となった。しかし、英語の各種標準テストの広がり、英語の重要性をグローバルに拡散することには貢献したが、現場で教師が担う英語教育を資格と標準スコアの結果に偏重したものへと変質させたと指摘されている (Cardoso & Steiner-Khamsi, 2017)。結果、独自のカリキュラムや教師が主体となって設計した現場レベルの英語教育の内容と成果は軽視され、標準スコアの結果のみが英語教育の成果として強調される風潮を強めた (Farhady, 2018)。

2. 研究の目的

本研究は、上記の背景を踏まえ、デザイン研究の蓄積を外国語教育に活用し、標準テストの特性を活かしながら英語の授業を構成する「実践」(Practice)・「評価」(Assessment)・「道具」(Tool & Technology) を有機的に統合する授業デザインの提案を目的とした (図 1)。

「評価」の要素では、標準テストの「標準化」と「比較可能性」の特性を活かしながらも、項目応答理論の制約(受験者数、項目数、テストの局所独立と次元性など)を克服し、小規模で且つ到達度の測定を原則とする日常的な英語の授業に対して適用可能な新たな数理モデルの提案を試みた。

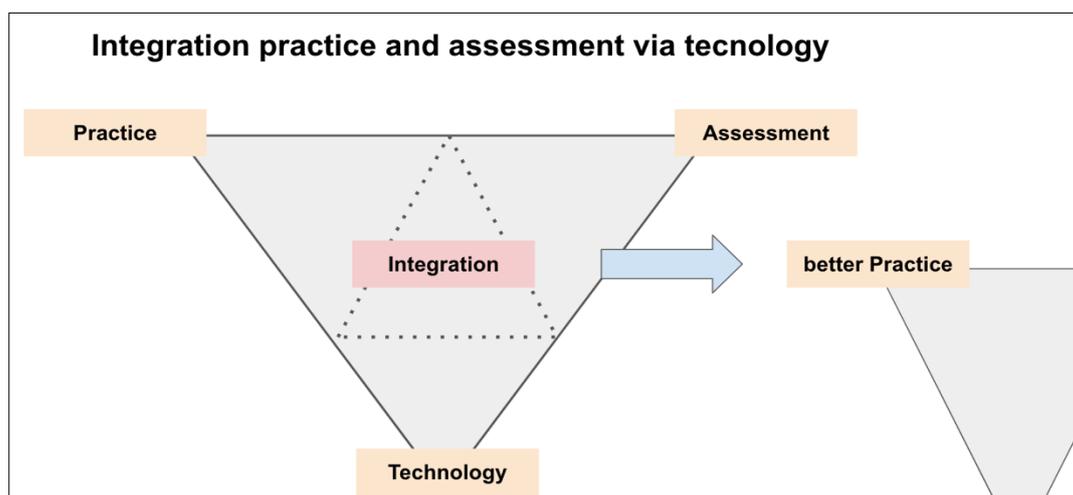


図 1 実践・評価・道具を統合した授業デザイン

「道具」の要素では、実践と評価を統合し、授業を伴奏しながら、さらに授業支援に貢献できる学習支援システムの新たな開発を目的とした。

以下、(1)「授業デザイン」、(2)「評価」、(3)「道具」、(4)「実践」の各要素について、具体的な研究方法を述べる。

3. 研究の方法

(1) デザイン

デザイン研究の省察プロセスを参考に、英語の授業の観察を複数回実施した。その結果、英語の授業には、「授業内フロー」(teaching flow within a lesson)と「授業間フロー」(teaching flow between lessons)が共通して存在することが分かった(図2)。教員間で微妙な差は存在するものの、授業内のフローでは、前回の授業内容の復習および各回の授業内容を達成するための授業が実施されていた。授業間フローでは、前回の授業内容を踏まえながらもシラバスに記載された授業目的に向けて授業内容が連動して蓄積されていくことが確認された。以上のことから、授業

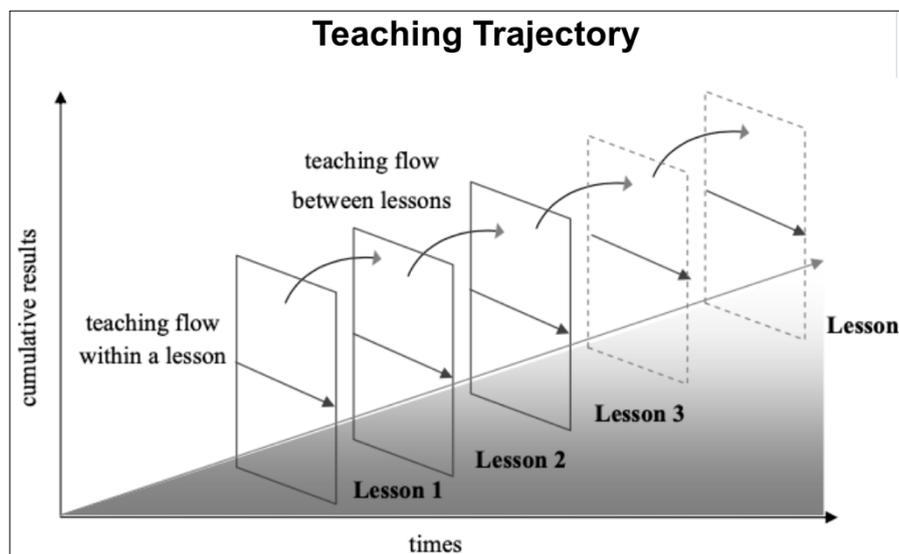


図2 授業内フローと授業間フローの関係性

内・間の連動性を阻害することなく支援する授業デザイン(実践・評価・道具の連動)とシステム開発の必要性が確認された。

(2) 評価

項目応答理論は、テスト結果の「標準化」と「比較可能性」という特性を持つ有益な数理モデルである。テスト市場に存在する標準テストのほぼすべては項目応答理論によって設計されているといっても過言ではない。しかし、その適用には厳しい理論上の条件が存在する。この条件のために外国語教育の実践や到達度の確認を目的とした小規模クラスの授業への適用例はほとんど存在しないのが現状であった。結果、外国語教育のような小規模なクラスでは標準スコアを活用して授業の成果を表現することが難しいのが現状であった。同時に標準テストの影響力は増し、一方で独自のシラバスに基づく教師を主体とした授業の成果は軽視される傾向にあった。この問題を解決するために複数の数理モデルを検討した結果、本研究ではイロレーティングを活用することにした。イロレーティングは、Elo (1978) によって提案されたもので、最近ではオンラインゲームやスポーツの勝敗に基づく個人またはチームのランキングの他、大規模テストの受験者能力推定にも用いられている (Brinkhuis & Maris, 2009; Brinkhuis & Maris, 2010; Papousek et al., 2014; Pelánek 2016)。

(3) 道具

道具の開発は、授業内・間フローを阻害せず実践と評価を連動させるシステムを検討した。また、イロレーティングを活用して、テスト結果から標準スコアを返す機能の実装を検討した。その結果、自動問題生成システム (Automated Test Maker、以下 ATM、<https://atm.lang-tech.net/>) を開発した。ATM は、①多肢選択問題、②読解問題、③適応型テストを自動的に生成することができる(図3)。多肢選択問題に対応した問題は、2,000 問以上、読解問題に対応したテスト冊子は約 80 種以上がアイテムバンクに格納されている。アイテムバンクの困難度パラメーターは、受験者の反応データが得られるとイロレーティングにより推定され更新される。適応型テストは、更新された困難度パラメーターに基づき、受験者の正誤反から受験者能力に適した問題を自動的に出題することができる。

(4) 実践

上記の(1) デザイン、(2) 評価、(3) 道具の要素を統合した授業実践を行った。実践は、2022 年 4 月～7 月(2022 年度春学期)の期間、理系学部 1 年生 108 名を対象に、英語読解の 4 つのクラスを対象に行った。多肢選択問題に関しては、4 月(春学期開始時)に 20 問からなる事前テ

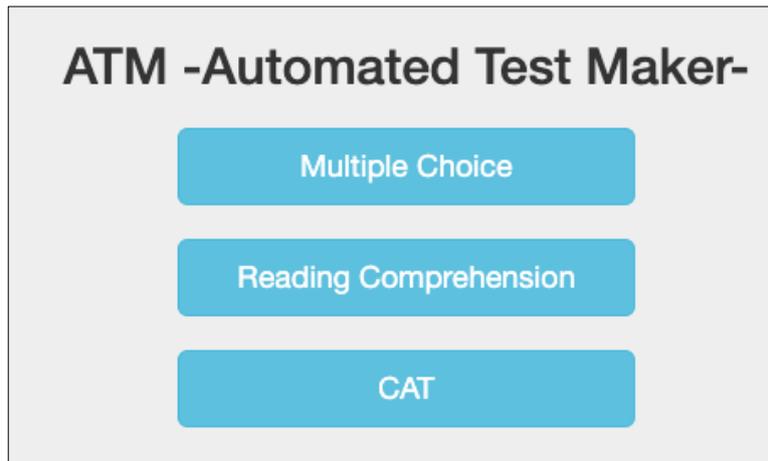


図 3 ATM のインターフェース

ストを実施した。春学期の期間中、語彙や解答方略のトレーニングを毎回の授業で行った。7月（春学期終了時）には、事前テストと同じ20問からなる事後テストを実施した。読解問題に関しては、30問からなる4種類（test A～D）のテストを作成し、4クラスを対象に7月（春学期終了時）に期末テストを実施した。30問中15問は4種類（test A～D）を横断する共通項目を含み、15問はクラス毎に異なる問題を出題した。加えて、実践の効果を多角的に検証するために、同じ英語読解の授業を担当する他のクラスの専任教員1名、非常勤講師3名を対象に半構造面接を行った。

4. 研究成果

多肢選択問題の事前テスト（ $n = 83, M = 13.23, SD = 3.17, \alpha = .638$ ）と事後テスト（ $n = 94, M = 17.67, SD = 2.71, \alpha = .763$ ）の結果を対応のある t 検定で分析した。その結果、有意な差を確認することができた（ $n = 81, t = -14.246, p = 0.000, d[95\% \text{ CI}] = 1.583[1.252, 1.908]$ ）。このことから、事前テストで学期のはじめに受験者能力値と、各受講生に共通する語彙や文法上の弱点を把握し、それを授業設計に活かし、毎回の授業で語彙や解答方略トレーニングを提供することの重要性を確認することができた。

読解問題に関しては、各テスト（test A～D）の問題項目をラッシュモデルとイロレーティングで推定した後に共通項目を使って等化し、相関係数の検討を加えた。その結果、test A は、 $r = .712$ （95% CI [.473, .853]）、test B は、 $r = .534$ （95% CI [.215, .75]）、test C は、 $r = .712$ （95% CI [.473, .853]）、test D は、 $r = .61$ （95% CI [.32, .796]）となった。この結果から、ラッシュモデルとイロレーティングの間に高い相関性を確認することができ、イロレーティングの有効性が確認された。

インタビュー調査では、「ATM は授業に役に立ちましたか？」という質問に対して、「毎回の授業で単語テストや小テストを作成することが大幅に楽になりました」、「理解の足りていない部分をすぐに把握できるようになったので助かりました」、「以前は問題作成から採点まで沢山の時間がかかっていましたが、今では一瞬でそれができるようになりました」といった意見が聞かれた。「ATM についてどう思いますか？」という質問に対しては、「LMS に比べて大幅に使いやすいです」、「授業を抜本的に変えてくれました」、「オンラインテストがとても便利でした」といった意見が聞かれた。この結果から、ATM の利用が学生の英語力の向上に寄与しただけではなく、授業担当者の負荷を軽減し、授業スタイルを変更することなく普段の授業で使える道具として有効に機能されていたことが分かった。

以上のことから、イロレーティングを実装した ATM を活用した授業デザインが、授業を質的に改善させ、学生の英語力の向上にも寄与していたことを確認することができた。また、イロレーティングが既存の項目応答理論と同様の推定精度を持ち、外国語教育のような小規模で到達度テストを原則とする授業スタイルにも応用可能であることを確認することができた。今後は、テスト項目の改善とアイテムバンクの拡充を進め、さらに推定精度を高めていくことが期待される。さらに、本研究によって開発された「実践」・「評価」・「道具」を有機的に連動させる授業モデルを他大学や中学・高校などの教育現場でも応用し、授業モデルのさらなる改善と ATM の機能向上に努めていくことが期待される。

<引用文献>

- Bakker, A. (2018). *Design research in education: a practical guide for early career researchers*. Routledge.
- Brinkhuis, Matthieu J. S., & Maris, G. (2009). *Dynamic parameter estimation in student monitoring systems. measurement and research department reports*. <https://www.researchgate.net/publication/242357963>
- Brinkhuis, Matthieu J. S., & Maris, G. (2010). *Adaptive estimation: How to hit a moving target*. Measurement and Research Department Reports.

- Cardoso, M., & Steiner-Khamsi, G. (2017). The making of comparability: Education indicator research from Jullien de Paris to the 2030 sustainable development goals. *Compare: A Journal of Comparative and International Education*, 47(3), 388–405. <https://doi.org/10.1080/03057925.2017.1302318>
- Elo, A. (1978). The rating of chess players, past and present. *Arco Pub*. <http://www.worldcat.org/title/rating-of-chessplayers-past-and-present/oclc/4504131>
- Farhady, H. (2018). *History of language testing and assessment* (J. I. Liontas (ed.); 1st Edition, pp. 1–7). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118784235.eelt0343>
- Goundar, P. R. (2020). An overview of language testing. *European Journal of Literature, Language and Linguistics Studies*, 4(3), 161–165. <https://doi.org/10.46827/ejll.v4i3.230>
- Himelfarb, I. (2019). A primer on standardized testing: History, measurement, classical test theory, item response theory, and equating. *Journal of Chiropractic Education*, 33(2), 151–163. <https://doi.org/10.7899/JCE-18-22>
- Papousek, J., Pelánek, R., & Stanislav, V. (2014). Adaptive practice of facts in domains with varied prior knowledge. *Educational Data Mining 2014*. https://educationaldatamining.org/EDM2014/uploads/procs2014/long%20papers/6_EDM-2014-Full.pdf
- Pelánek, R. (2016). Applications of the Elo rating system in adaptive educational systems. *Computers & Education*, 98, 169–179. <https://doi.org/10.1016/j.compedu.2016.03.017>
- van den Akker, J. J. H. (Ed.). (2006). *Educational design research*. Routledge.
- Vinovskis, M. A. (2019). History of testing in the United States: PK–12 Education. *The Annals of the American Academy of Political and Social Science*, 683(1), 22–37. <https://doi.org/10.1177/0002716219839682>

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 住政二郎・光永悠彦・出口朋美
2. 発表標題 イロレーティングを応用した Automated Test Maker の開発
3. 学会等名 日本教育工学会 2021年秋季全国大会
4. 発表年 2021年

1. 発表者名 Sumi, S., & Deguchi, T.
2. 発表標題 Application of Elo Rating System to Small-Class-Sized Achievement Tests
3. 学会等名 56th RELC International Conference (国際学会)
4. 発表年 2022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	光永 悠彦 (Mitsunaga Haruhiko) (70742295)	名古屋大学・教育発達科学研究科・准教授 (13901)	
研究分担者	出口 朋美 (Deguchi Tomomi) (90708538)	近畿大学・法学部・准教授 (34419)	

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0件

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
---------	---------