

令和 元年 9 月 10 日現在

機関番号：82616

研究種目：国際共同研究加速基金（国際共同研究強化）

研究期間：2017～2018

課題番号：16KK0010

研究課題名（和文）情報端末・統計理論・自然言語処理技術を用いた先端的試験技術の開発と運用（国際共同研究強化）

研究課題名（英文）Advanced testing technology based on information device, statistical theory, and natural language processing(Fostering Joint International Research)

研究代表者

大久保 智哉（OKUBO, Tomoya）

独立行政法人大学入試センター・研究開発部・准教授

研究者番号：80512136

交付決定額（研究期間全体）：（直接経費） 10,000,000円

渡航期間： 16ヶ月

研究成果の概要（和文）：本研究において、試験結果を分析・報告可能なシステムの研究開発、海外における先端的試験技術の利用最新動向に関する調査および検討、従来の教科科目型ではない能力特性のための測定方法に関する調査研究、先端的試験技術の技術的発展、に関して成果が得られた。特に、従来よりも解答時に作業を要する試験課題を用いた非教科教育型のコンピュータ上で実施する試験問題については、システム上の困難さよりも試験運用上の具体的な課題そのもの、および試験問題の方策について、本研究遂行時に国際共同研究によって多くの示唆を得ることができた。

研究成果の学術的意義や社会的意義

本研究プロジェクトは、英国ケンブリッジ大学に加えて仏国の国際機関である経済協力開発機構の教育局において遂行されたが、これらの機関では試験実施デバイスについて紙筆型の試験からコンピュータ端末による実施へと転換している。そして、世界各国において、彼らによって実際に運営される試験システム・運用が稼働しており、単に研究上の進展だけでなく、研究テーマとなっている先端的試験技術について実際の試験オペレーションについて検討を重ねることができた。このことは、我が国において進められている高大接続改革への還流が大いに期待されるものである。

研究成果の概要（英文）：This research project accomplished the following four points: 1) Development of the test analysis system that enables us to report the test results visually and dynamically. 2) Research on usage of the testing technology abroad, especially in the United States, the United Kingdom, the Netherlands, and international organisation such as the OECD (Organisation for Economic Co-operation and Development). 3) Research on item development in non-subject based tests, such as innovative domain. Concretely, creative thinking and global competences. 4) Development in the advanced testing technology based on computer-based testing. Especially, we focused on the management of the non-subject based testing and test item itself.

研究分野：教育測定学

キーワード：試験技術

1. 研究開始当初の背景

大学入学試験や全国学力・学習状況調査などの大規模テストは、学習者の学習方略や教育者の教育内容に大きな影響を及ぼすことが明らかになっている。大学入試が変わるとそれに伴って高校・中学・小学校の順に教育内容が変わっていく。しかし、現在の試験には実施面において限界があり、効果的な試験を実施できていない。「大学入試が変わらないから学校教育が変わらない」という多くの教育関係者の声はその現状を表している。現在の試験の限界は、以下の2点に起因している。1) 異なる試験実施回での得点比較ができないこと：そのため、厳密な得点比較が求められる入学試験などは「年に一度の試験の一斉実施」という強い制約が課されることとなる。この制約が「同一試験の点数による受験生の序列化」を促し、結果「年に一度の試験で1点でも多く取る」ことを主とするような教育が増えている。2) 大規模試験ではマークシートを採用することになること：大規模試験は採点時間と厳密性およびコストの観点から、マークシートによる多肢選択形式にならざるを得ないという制約も生み出している。さらに一点を争う試験では試験問題に対しても厳密性が求められるため、曖昧さを排除された知識を問うような問題にならざるを得ない。

このような制約が、実際に社会や教育現場で求められる能力とは異なる能力を測定することを促し、教育の発展を阻害している側面がある。試験運用上の制約をクリアすることを通じて試験内容の改善を図ることは、教育内容の改善にも資する重要な研究課題となっている。基課題においては、十分な研究成果と大きな社会へのインパクトを与えてきたが、本研究の成果をさらに発展させるためには、国際共同研究において以下の点を達成する必要がある。1) 能動技能の採点技術の研究開発：現在、社会では speaking 試験や writing 試験などの能動技能の測定が求められている。基課題においては、システムの実験的運用をおこなっているが、英語4技能試験などの円滑な多技能試験導入のためには、効率的な採点技術の整備が不可欠である。2) 思考力・判断力・表現力の評価モデル構築への世界的潮流の加速：すでに、PISA などの国際学力調査では、思考力や判断力を問う新たな試験内容を情報端末上で試行している。また、国際共同研究先の John Rust 教授の研究チームは、オンライン上で試験を実施可能にするオープンソース(Concerto)を公開して、様々な認知能力の測定を可能にする新たな試験に関する研究についても世界的に主導している。

2. 研究の目的

本研究の目的は、

- 1) 能動的技能に関する試験問題およびその採点技術の研究開発。
 - 2) コンピュータ上で可能となる思考力や判断力を問うような試験の研究開発。
 - 3) コンピュータ利用型試験の世界の研究潮流の調査およびそれらを可能にするシステムの研究開発
- の3点である。これらの研究は基課題をもとに発展させる形となる。

3. 研究の方法

上記研究目的を達成するために、英国のケンブリッジ大学(University of Cambridge, UK)を主な国際共同研究先とした。具体的には、Cambridge 大学の John Rust 教授のチームと基課題において開発された試験技術および採点技術の共有、さらには信頼の元でお互いのシステムについての改善点の議論および改善をおこなった。このことは、開発されたシステムや技術の安定化や発展に寄与した。また、ケンブリッジ大学、テスト開発チームで進められている思考力・判断力・表現力の評価モデルに関する研究に参画することにより、能動的な能力特性を測定する試験について検討をおこなった。さらに、英語の能動技能の評価について、同大学の研究者らと共同研究を通じて世界の潮流の把握につとめた。さらに、仏国 OECD 教育局において、創造的思考やその他能動的な能力特性を測定する方法について、研究を進め実際に世界のいくつかの国から実際に試験をおこなった上でデータを取得し分析をおこなった。仏国 OECD では、世界80カ国で実施されている PISA 調査をおこなうチームとともに調査研究をおこなったことがシステムや測定する能力特性に関する世界の潮流などの調査研究につながった。さらにその結果に基づいて、さらなる研究をすすめた。

4. 研究成果

本研究において、コンピュータ利用型試験システムの研究開発および海外における先端的試験技術の利用最新動向に関する調査および検討、加えて、従来の教科科目型ではない能力特性のための測定方法に関する調査研究とその採点のための先端的試験技術の研究、について研究成果が得られた。特に、従来よりも解答時に作業を要する試験課題を用いた非教科教育型

のコンピュータ上で実施する試験問題については、実際に共同研究者らと開発をおこない、検証のためのデータを取得して結果を分析した。さらに、それらを作成するに必要なシステムについて検討を加えた。なお、研究成果については国際学会における招待講演および発表として公表されている。

研究成果として、第一に、先端的試験技術の利用最新動向に関する調査および検討をおこない知見を得た。特に、コンピュータ利用型試験が世界においてどのような技術をもって、どのように実際に展開されているか、という点について調査をおこない検討をした。具体的には、大規模試験および調査を対象にして、展開している試験実施システムの機能、試験問題における解答形式、試験実施時の運営手順、試験問題の作成方法に加え、それらを統合的に取り扱う試験運用システムについてその機能要件について調査研究をおこなった。さらに具体的にその研究成果の一部については、大学入試センター試験に代わる新試験である大学入学共通テストに導入することが検討されているコンピュータ利用型試験システムの機能として導入をおこなった。

次に、従来に日本で多く用いられてきたような教科型ではない、非教科型の試験問題について、共同で研究開発し、実際に複数国に対して検証のための調査をおこないデータを取得した上で、分析してその試験問題の特性について検討をおこなった。具体的には、創造的思考力などを対象にした試験問題であった。またその際に適切な試験解答形式についても検討を加えた。そして、最後に思考力・判断力・表現力に対する調査方法に関して世界の動向について調査検討をおこなった。また、大規模調査において必ず生じる受験者集団の特性による試験問題の特異的機能について、その検出方法および分析方法について調査研究をおこない、シミュレーションをおこなうことで最も適切な分析モデルについて検討をおこなった。このような特に大規模試験・調査について特に発生しうる問題について調査研究をおこなった。

さらに、共同研究者のプロジェクトにおいて、創造的思考(Creative Thinking)に関して、コンピュータ上で実施可能な試験問題およびその採点基準を作成した。さらに、その試験問題についてオーストラリア、シンガポール、南アフリカ共和国、カナダにおいて、検証用の調査をおこないデータを得た。さらに統計的分析をおこなうことで、試験問題および採点基準を評価し、さらなる改善につなげることができた。これらは能動的態度についても測定対象とするような試験問題であった。さらに、上記のような新しい能力概念観に基づく試験問題に対して、自然言語処理技術を援用して効率的に採点をおこなう手法について研究をおこなった。研究成果については、国際学会において発表されている。

本研究プロジェクトは、英国ケンブリッジ大学に加えて仏国の国際機関である経済協力開発機構の教育局において遂行されたが、これらの機関では試験実施デバイスについて紙筆型の試験からコンピュータ端末による実施へと転換している。そして、世界各国において、彼らによって実際に運営される試験システム・運用が稼働しており、単に研究上の進展だけでなく、研究テーマとなっている先端的試験技術について実際の試験オペレーションについても調査研究を重ねた。

5. 主な発表論文等 (研究代表者は下線)

〔雑誌論文〕(計 件)

〔学会発表〕(計 2件)

1. Okubo, T. (2018). New common university entrance examinations with testing technology. *The 2018 Global Summit on Artificial Intelligence and Big Data in Education*. Beijing.
2. Primi, R., Piacentini, M., & Okubo, T. (2019). Exploring the use of deep learning to score tasks of PISA 2021 assessment of creative thinking. *International Association for Computerized Adaptive Testing 2019*, Minneapolis.

〔図書〕(計 件)

〔産業財産権〕
出願状況(計 件)

名称：
発明者：
権利者：
種類：
番号：

出願年：
国内外の別：

取得状況（計 件）

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織

研究協力者

〔主たる渡航先の主たる海外共同研究者〕

研究協力者氏名：John Rust

ローマ字氏名：

所属研究機関名：University of Cambridge

部局名：The Psychometrics Centre

職名：Professor

〔その他の研究協力者〕

研究協力者氏名：

ローマ字氏名：

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。