

令和 5 年 6 月 12 日現在

機関番号：14301

研究種目：挑戦的研究（萌芽）

研究期間：2017～2022

課題番号：17K18627

研究課題名（和文）論述試験の知見を基盤とした適応型面接試験の開発による全人的教育測定法の探索的研究

研究課題名（英文）An exploratory research on holistic educational measurement with adaptive interview testing based on previous writing assessment findings.

研究代表者

大澤 公一（Osawa, Koichi）

京都大学・高大接続・入試センター・特定准教授

研究者番号：20555320

交付決定額（研究期間全体）：（直接経費） 4,800,000円

研究成果の概要（和文）：言語4技能を中心とした日本語運用能力のほか、大学をはじめとする高等教育機関における学修活動で必要となる様々な能力（学力）や適性・資質等を、口述能力を媒介して測定・評価を行うための統合的な試験の試作を研究目的とした。試験のプラットフォームはiPadを用いた適応型web面接試験とし、総合的な言語能力に加えて一般知識、論理的思考能力、問題解決能力、表現力、コミュニケーション能力などを統合的に評価する対話型試験の作成を目指した。当初は海外留学生の獲得促進（大学の国際化）を目的とした外国語母語話者の日本語学習者をターゲットとしたが、コロナ禍により海外モニター試験の実施は叶わなかった。

研究成果の学術的意義や社会的意義

従前の心理学的あるいは教育学的な能力測定においては、ある能力を測定・評価するためには当該能力に対応した個別の試験を（従って複数能力の場合には能力の個数分）実施することが大前提となっている。本研究は、挑戦的研究（萌芽）として、学力等の能力・適性や言語能力の諸技能の測定を、口述能力を媒介した適応型の面接試験によってある程度統合することにより、従前の能力別の試験の開発や実施の負担軽減を試みたものである。一方で、大学などの高等教育機関への志願者の教育背景が多様化するに伴い、従前の学力検査に偏重した入学者選抜方法からの脱却を目指し、総合的・統合的な能力評価の方法を模索するという意義もある。

研究成果の概要（英文）：This research was planned to develop an integrative testing platform to measure and evaluate, via speaking competency, academic abilities, psychological traits and aptitudes as well as Japanese language abilities based on CEFR 4 language skills, required when studying in Japanese undergraduate course. An adaptive web interview test utilizing iPad was introduced as a testing platform, whose content included general knowledge, logical thinking, problem solving, productive expression, communication skills as well as comprehensive language abilities. At the beginning years of the research, Japanese language learners in overseas were supposed to be considered as the main subjects of the research. This is because there have been seen a trend in Japanese universities to recruit international student from overseas. However, unfortunately, COVID-19 prevented the researcher from implementing monitor testing in overseas.

研究分野：教育心理学

キーワード：教育心理学 テスト理論 教育測定 心理測定

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

(1) 社会的背景

現在の大学入学者選抜制度の改革において、学力は「知識・技能」、「思考力・判断力・表現力」、「主体性・多様性・協働性」の3要素に分類される。受験者の能力や適性・資質等の多面的・総合的評価(全人的教育評価)が求められるながら、現行の入試制度では新しい学力観を適切に測定・評価することは難しく、論文や面接試験などの測定技術の再開発が必須である。

・言語能力は4技能(読解, 聴解, 論述, 口述)で構成されると考えられており、各技能を測定する試験がアラカルトで提供されるのが標準である。しかしながら、大規模な試験においては論述や口述試験の実施が困難なため、言語能力を十分に測定・評価できないケースがある。

・項目反応理論(以下IRT)やテスト等化法に代表される現代テスト理論が実用の段階に入った。今後は、コンピュータ適応型テストやICTを基盤とするwebテストなどの実装が進み、それに合わせてテストの内容や構成も一新されることが期待される。

(2) 学術的背景

文部科学省(大澤, 2007~2008年度)および科学研究費補助金(大澤, 2010~2012年度; 2013~2016年度)の助成を受け、以下~の研究が行われた(図1)。

大規模公的日本語試験を客観的な参照基準とした新しい日本語論述試験の開発  
 論述試験の項目反応理論による尺度化・得点等化法の開発  
 日本語論述能力の縦断的・診断的評価による連続測定法の開発  
 高等教育において必要となる適性・資質、スキルや言語運用能力4技能に関するCan-do Statements 尺度との相関関係を明らかにするための質問紙調査  
 既存の大規模公的試験の既出問題を用いたモニター試験

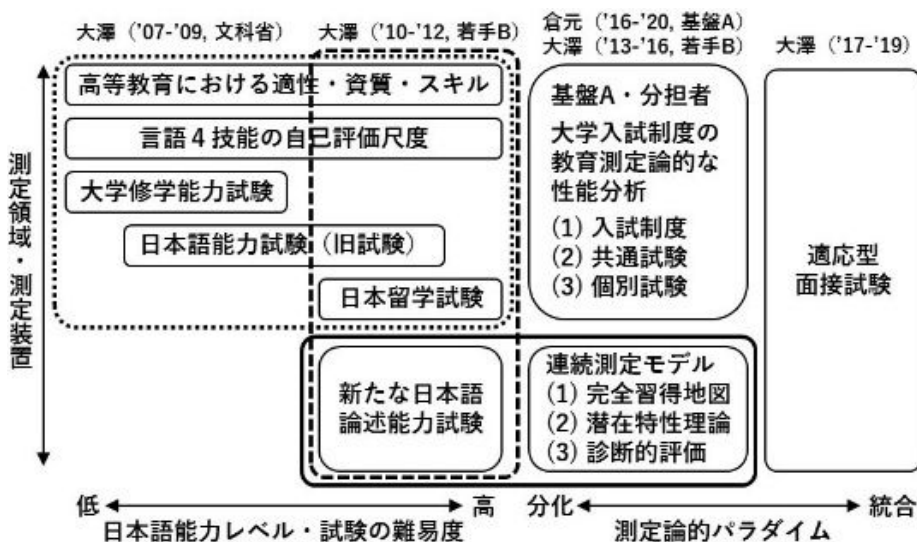


図1. これまでの研究の流れ

(3) 着想に至った経緯

申請者は外国語母語話者を対象とした日本語能力の測定法を研究テーマの一つの柱とし、適性試験型の日本語論述課題の研究開発に取り組んだ。これを基に、日本人学生に対する日本語論述能力の連続測定モデルに基づく診断的教育測定エンジンの基礎的な研究に取り組み、論述能力の測定法の精緻化と測定評価領域の拡張・学習カリキュラムとの連続的接続に挑戦している(図1)。

一連の研究の流れを踏まえると、現状では言語4技能の測定は各技能に対応するテスト(読解・聴解・論述・口述)に依存している。機能分化された測定装置の一つとして論述能力の測定法の研究は継続するが(基盤A・分担者)、学力等の能力・適性や言語の諸技能を統合的な観点から測定・評価する研究が行われていないため、本研究の構想を立ち上げた。

## 2. 研究の目的

前頁の(1)社会的背景および(2)学術的背景を踏まえ、申請者の10年間に渡る先行研究によって収集された膨大かつ客観的なデータ(エビデンス)に基づいて、日本語論述能力試験に関する知見を基礎とした適応型面接試験を開発することが本研究の目的である。日本語運用能力だけでなく、高等教育において必要となる受験者の様々な能力(学力)や性格・適性・資質等を測定対象とし、対話型の口述試験において統合的に測定・評価するための新しい試験を開発する。

適応型面接試験では、質問項目に対する受験者の能力レベルが反映された個別の反応に応じて、次に提示する質問内容が変化していく。測定される能力や性格・資質には、知識、言語運用能力、論理的思考能力、問題解決能力、即時判断力、表現能力、コミュニケーション能力、性格特性、社会的適応力などが含まれる。これらの潜在特性を、分析的ないしは総合的な評価基準によって尺度化し、受験者の能力プロファイルを作成する。

本研究における適応型面接試験は大学入試などの選抜場面での使用を前提とし、診断的評価を第一義的な目的とはしない。試験はwebを通してオンライン上で行うことを前提とし、そのため試験プラットフォームを標準的に構成・提供するものとする。

## 3. 研究の方法

研究助成期間を通しての下位研究の当初の計画は以下の通りであった。本研究の当初計画においては、海外の研究協力者の助力を得た上で日本語学習者に対するモニター試験の実施を予定していた。しかしながら、2019年末からの新型コロナウイルス感染症の世界的流行により、当初の研究計画にあった海外モニター試験の実施が困難となり、それに伴い研究計画の大幅な変更を余儀なくされることとなった。

コロナ禍に入った後、令和2年度~4年度までの期間を研究(特別)延長期間として申請し、状況の好転を待ちながら研究計画の実現可能性を慎重に検討した。日本においては、新型コロナウイルス感染症は2023年5月8日をもって感染症法上の位置づけが5類感染症に変更されるなど、状況の相対的な好転を見せている。しかしながら、諸々の状況を総合的かつ慎重に検討した結果、当初に計画していた海外モニター試験の実施は見送らざるを得ないと判断し、令和4年度末に国内においてiPadをプラットフォームとする試行試験を行った。

### (1) 平成29年度(2017年度)

口述能力を媒介して言語4技能や学力等の測定を行うための課題の開発

... 先行研究(大澤, 2010~2012)に基づき、新型論述試験課題の深化・発展

... 学力レベルは大学入試センター試験を越えないものとする

能力(学力)を総合的に尺度化するための評価基準の開発

... 先行研究(大澤, 2013~2016)に基づき、分析的評価尺度と総合評価尺度の精査

口述回答用の適性検査の開発

... 先行研究(大澤, 2007~2016)に基づき、評価項目の精査

... 主要大学の入試追跡調査研究資料の収集と要約

... 豪州、英国のメディカルスクール入試で用いられる適性検査の収集と翻訳

海外の高等教育機関における入学試験(面接試験)の詳細な構造分析

... 英国の大学入試における面接試験の内容に関する調査

### (2) 平成30年度(2018年度)

続・口述能力を媒介して言語4技能や学力等の測定を行うための課題の開発

... 各教科につき、数段階のステップを経ないと正解に至らない課題の作成・体系化

... 正解へのステップごとに、受験者の能力水準に応じたヒント(枝問題)の設定

続・能力(学力)を総合的に尺度化するための評価基準の開発

... 総合得点の分散を最大とするような尺度化のための数理統計学的研究

続・口述回答用の性格・適性検査の開発

... 主要評価項目の確定

先行研究と紐付けするためのモニター試験・質問紙調査データの解析

... 先行研究(大澤, 2007~2012)に基づき、客観試験と適性自己評価IRT尺度の構成

### (3) 平成31年度(2019年度), 令和2年度~4年度(2020~2022年度, 延長期間)

面接試験(対話型試験)を行うためのwebプラットフォームの開発

... 適応型面接試験は留学生の獲得促進(大学の国際化)を主目的の一つとする

... オンライン(web)での実施を想定

... 中国, 香港, 台湾, 韓国, インドネシア等でのモニター試験の準備(実施断念)

モニター試験の実施およびデータ解析

... データ分析における統計学的な技法は先行研究(大澤, 2007~2016)に準拠

#### 4. 研究成果

本研究課題の成果に関して、その最も顕著なものは、言語運用能力の構造がどのような潜在特性（構成概念、因子）から構成されているのかを、大規模なデータから測定論的に明らかにしたところであろう。本研究の当初の目的であった、外国語母語話者の（外国語による）口述能力を媒介した教育測定を行うためには、第一外国語の言語運用能力の潜在構造を明らかにする必要がある。そこで、韓国語を母語とする日本語学習者、ならびに日本語を母語とする英語学習者に対してそれぞれの学習言語（日本語あるいは英語）の言語運用能力を自己評価するための尺度を構成し、その結果に対して因子分析を行った。以下にその概要を紹介する。

##### (1) 韓国語を母語とする日本語学習者の分析（日本語運用能力）

JSL 学習者を対象として日本語運用能力 4 技能（読む、書く、聴く、話す）の自己評定による保有の程度、および高等教育で必要とされる適性や資質・スキルとの相関関係に基づく潜在構造を探索的因子分析により明らかにし、留学生入試や JSL 学習者に対する大学教育の改善に資する基礎データを構成することが本分析の目的である。

韓国内の 115 校の教育機関（高等学校、短期大学校、大学校など）より韓国語母語話者の JSL 学習者 4,576 名の協力を得た。分析対象とする尺度項目は、日本語 Can-do-statements 尺度より 4 技能 60 項目（各技能 15 項目）、高等教育で求められる適性や資質（27 項目）・スキル（20 項目）の自己保有度に関する調査より 47 項目、合計 107 項目であった。

標本数  $N=4,576$  のうち欠測値を含む観測値は  $N=945$  であり、全回答データの 20.65% に相当した。このデータセットに対して欠測値を含む 945 のデータを削除した上で因子分析を行う方法は合理的ではない。本研究では、多重代入法により 40 セットの疑似完全データセットを生成し、不完全データセットの共分散行列の推定を行った。推定された共分散行列（相関行列）に対して平行分析を行うと、推奨抽出因子数は 13 であった。次に Very Simple Structure 基準（MAP）で分析を行うと、推奨抽出因子数は 11 であった。そこで、抽出因子数の目安を 11～13 とし、探索的因子分析（最尤法・オブリンミン回転）を行った。

平行分析によると推奨抽出因子数は 13 であったが、ある単独の因子について 0.3 以上の因子負荷をもつ項目が存在せず、因子の命名および解釈が困難であった。13 因子モデルと 12 因子モデルおよび VSS 基準に基づく 11 因子モデルとの違いは、様々な領域に対する関心において審美・芸術への関心がその他の関心と分離されるかどうかという一点であった。本研究のテーマにおいて両者を積極的に分離する理由が存在しないため、「対象領域への関心」を単独の因子として測定することとし、結果的に 11 因子モデルを採用した（表 1）。

日本語 Can-do-statements 尺度に関しては、4 技能に直接相当する因子として「読む」「書く」「聴く + 話す」という 3 因子に分割された。また「上級の言語運用能力」および「学校生活（特定の目標言語使用領域）」という命名・解釈が可能な 2 因子が抽出された。本研究で音素に係る 2 つの技能（聴く、話す）が単独因子に収斂した結果は、CEFR におけるモードに基づく言語運用能力モデルの考え方と整合性が取れていた。なお、日本語能力および適性・スキルが同時に負荷するような潜在因子は、本分析において抽出されなかった。

高等教育で求められる適性や資質・スキルに関しては「一般適性」「数理スキル」「論理的思考力」「言語運用能力」「ICT スキル」「対象領域への関心」という 6 因子が抽出された。なお、探索的因子分析を完全データ（ $N=3,631$ ）で行った場合でも日本語能力、適性スキル両者において多重代入法に基づく本分析手法と類似の因子パターンが得られるものの、各項目の因子負荷の状況は、多重代入法に基づく正確な分析結果と完全には一致しなかった。

表 1. 因子間相関（韓国語を母語とする日本語学習者）

	読	書	聴話	上級	学校	適性	数理	思考	運用	ICT	関心
読	1										
書	.745	1									
聴話	.774	.702	1								
上級	.381	.257	.179	1							
学校	.565	.582	.648	.145	1						
適性	.191	.200	.234	.014	.164	1					
数理	-.088	-.077	-.094	.013	-.130	.245	1				
思考	.102	.094	.103	.068	.043	.477	.451	1			
運用	.169	.176	.217	-.051	.151	.469	.217	.494	1		
ICT	.215	.185	.223	.032	.113	.308	.420	.393	.221	1	
関心	.258	.215	.243	.106	.158	.322	.175	.328	.183	.321	1

(2) 日本語を母語とする英語学習者の分析（英語運用能力）

英語運用能力の潜在構造を明らかにするために、ある日本国内の大学の学部新入生を対象として CEFR に基づいた言語運用能力の自己評定質問紙調査を 4 年度に渡って実施した。本研究で用いられた 114 の質問項目群は、年度間で共通の項目群が配置される共通項目デザインによって各年度の調査対象者に提示された。調査は平成 30 年度および平成 31 年度が紙筆版、令和 2 年度および令和 3 年度はコロナ禍のため web 版で行われた。調査対象者は、提示された能力記述文ごとに、自分自身の英語運用能力の程度を教示に従って 4 段階で自己評定を行った。

調査への回答が得られたのは累計 8,023 名であった。平成 30 年度の回答者は 2,398 名（回収率 80.99 %）、平成 31 年度は 2,473 名（同 81.34 %）、令和 2 年度は 1,493 名（同 54.91 %）、令和 3 年度は 1,659 名（同 61.10 %）の累計 8,023 名（同 70.73 %）であった。

上述の 4.(1)と同様に 8,023 名×114 項目のデータ行列は欠測値を含む不完全データであり、観測データに対して直接的に因子分析を実施することには測定論的な問題がある。本分析でも、多重代入法により不完全データの共分散行列を推定した。推定された共分散行列に対して因子分析を実行することによって推定バイアスを回避し、因子パタンの一様性を担保した。Very Simple Structure 基準（MAP）による分析を行ったところ、抽出因子数は 12 であり、期待分散説明率は 0.679 であった。上記の分析を踏まえて因子分析（12 因子，最尤法，オブリミン回転）を実行した。探索的因子分析の結果，4 技能に相当する因子と共に，産出，相互作用，媒体などに関する因子群が抽出された。

因子分析で得られた 12 因子の因子間相関を表 2. に示した。CEFR の能力記述文を中心とした尺度項目の潜在構造は，言語 4 技能に相当する因子（読む：R1，R2，書く：W，聴く：L，話す：S1，S2，S3）に加えて CEFR における言語運用能力モデルを構成する概念（モード）に近い因子（産出：Pd，音声媒体：M1，表記媒体：M2，M3，技能総合：GL）が抽出され，研究代表者が 12 の因子群の命名を行った。

表 2. 因子間相関（日本語を母語とする英語学習者）

	R1	R2	W	L	S1	S2	S3	Pd	M1	M2	M3
R1	1										
R2	.453	1									
W	.415	.386	1								
L	.368	.368	.395	1							
S1	.285	.489	.380	.542	1						
S2	.275	.235	.409	.482	.487	1					
S3	.389	.293	.513	.334	.289	.364	1				
Pd	.190	.364	.320	.275	.545	.302	.242	1			
M1	.435	.273	.370	.269	.158	.243	.477	.186	1		
M2	.441	.400	.349	.387	.431	.387	.156	.278	.152	1	
M3	.353	.253	.415	.237	.135	.202	.455	.171	.372	.180	1
GL	.363	.260	.497	.254	.220	.294	.480	.184	.521	.171	.367

(1)(2)の両研究は，それぞれの研究目的が独立しているため，調査に用いた質問項目が同一のものではない。そのため，探索的因子分析の結果（抽出された因子）を直接比較することは不適切である。しかしながら，対象言語が異なれば，学習者の言語運用能力の潜在構造も異なる可能性を一定程度は示唆している。

本研究では，特に(1)の結果に基づいて，言語（日本語）運用能力と高等教育における適性・資質との間の因子間相関が比較的小さいことを重要視し，これらの能力を統合的に測定・評価することの合理性を担保するための積極的なエビデンスは得られなかった。本研究は，学力等の能力・適性や言語運用能力の測定を適応型面接（口述）試験の枠組みで統合する試みであったが，能力の測定においては，各能力に対応した個別の試験項目を用意する必要があるという前提を保持する必要があると推察された。

その一方で，言語運用能力と高等教育における適性・資質との相関が全くないわけではなく，種々の能力を測定・評価する上での言語能力の媒介性は存在すると考えることの合理的は残されている。その意味では，面接・口述試験（生産的言語運用能力）の枠組みで種々の能力評価を行うことに対する合理性までが否定されたわけではない。

本研究は挑戦的・探索的な研究であったが，研究期間中に作成されたテスト項目（非公開）を用いて，更なる実データ（解答テキスト・音声データ）の集積を行っていく必要がある。また，海外モニター試験を実施することで，遠隔実施可能な適応型面接試験としての基本性能を向上させていくことが期待される。それに加えて，対象言語（日本語・英語）の違いによる言語運用能力の潜在特性の構造の差異を考慮に入れた言語能力の評価・測定手法の検討を進める必要が明らかになった。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 大澤公一	4. 巻 30
2. 論文標題 自己評定型のCEFR英語運用能力IRT尺度を構成するための尺度候補項目の選定および評定反応データの基礎的分析	5. 発行年 2020年
3. 雑誌名 大学入試研究ジャーナル	6. 最初と最後の頁 132-139
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計4件（うち招待講演 1件 / うち国際学会 1件）

1. 発表者名 大澤公一
2. 発表標題 日本語 Can-do-statements 尺度および高等教育で必要とされる適性や資質・スキルの 自己評定尺度の相関関係に基づく潜在構造の探索的分析：韓国語母語話者の JSL 学習者を対象として
3. 学会等名 日本教育心理学会
4. 発表年 2021年

1. 発表者名 大澤公一
2. 発表標題 CEFRに基づく英語4技能運用能力に関する自己評定尺度の構成
3. 学会等名 日本テスト学会
4. 発表年 2019年

1. 発表者名 Koichi Osawa
2. 発表標題 A basic research for test equating and linking of College Scholastic Ability Test Japanese I
3. 学会等名 The 2018 International Conference on Japanese Language Education (国際学会)
4. 発表年 2018年

1. 発表者名 Koichi Osawa
2. 発表標題 Leading Directions of Educational Changes: Bringing global perspectives and local initiatives together
3. 学会等名 The KICE's 20th Anniversary International Seminar (招待講演)
4. 発表年 2018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
---------------------------	-----------------------	----

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