

論文内容の要旨

論文題目 DEVELOPMENT OF THE NUMERICAL REASONING ABILITY TEST AND
THE VOCABULARY COMPREHENSION ABILITY TEST
FOR MYANMAR HIGH SCHOOL STUDENTS

(ミャンマーの高校生のための数的推理能力テストと
語彙理解能力テストの開発)

氏 名 AYE AYE MYINT

【問題と目的】

現在ミャンマーの教育現場では、大学の入学選考などの重要な決定を行う際、学生の学業成績に加え、教科の学力以外の様々な能力、あるいは学内外での活動状況なども広く評価し活用している。しかし、こうした能力評価は教師の主観的判断に基づいて行われており、客観的な評価の枠組みはほとんど確立されていないのが現状である。一方、欧米諸国においては、能力評価を客観的に行うために、標準化された知能検査や学力検査が開発され利用されている。このような標準化されたテストの開発はミャンマーにおいても急務であると考えられる。能力評価の中でも特に重要視されているのは、数的推理能力と語彙理解能力であり、これらは教科学習とも深く関連する。そこで本研究では、ミャンマーの高校生のための数的推理能力テストおよび語彙理解能力テストを開発することを目的とする。

テスト開発にあたっては、近年、欧米各国において広く研究され、かつ利用されているテスト理論である項目反応理論 (Item Response Theory; IRT) を用いることにする。本研究では、IRT モデルのうち、2-パラメタ・ロジスティックモデルに着目し、活用することにする。IRT モデルを活用することの利点は、項目プールを構成できること、適応型テストを実施できること、項目バイアス

を検討できることなどであり、従来のテスト理論にはない多くの利点を有している。

また、現在ミャンマーでは、高校生に対する統一学力テストは実施されていないため、各学校ごとに実施されている月例教科テストの成績を基準とする能力テストの妥当性の検討を行う。その際、教科の成績が共通のテストで測定されていないことが問題となるが、本研究では、ベイズ統計学的手法である μ 集団回帰を応用することによりこの問題に対処する方法を考え、妥当性の検討を行う。

【数的推理能力テストの開発】

1. 項目作成：数的推理能力に関する3領域（文章題・数値方程式・数列）について、諸外国で開発されているいくつかの知能検査や適性検査を参考に、自由回答形式の項目を80項目作成した。それらについて10名の高等学校の数学科教師およびヤンゴン教育大学の教官の検討を受け、さらに、5名の高校生にすべての項目を実施し、全体の回答時間、各項目に対する回答速度、テスト後の感想などの情報を収集した。これらの情報をもとに、難易度が高すぎる、冗長であるなどと判断された24項目を削除し、一部の項目内容を修正して、56項目からなる予備テストを作成した。

2. 予備テスト：1996年7～8月に、180名の高校生を対象に予備テストを実施した。得られたデータをもとに、項目正答率の検討、因子分析による1次元性の確認、IRTモデルを適用した項目パラメタの推定などを行い、いくつかの項目を入れ替えて、45項目からなる本テストを作成した。

3. 本テスト：1996年9～10月に、多段階比例抽出法で抽出した905名の高校生標本を対象に本テストを実施した。得られたデータをもとに、項目分析、因子分析による1次元性の確認、因子負荷が0.35以下の項目の削除を行い、最終的に25項目からなる数的推理能力テスト（Numerical Reasoning Ability Test; NRT）を作成した。因子分析の結果をみると第1因子の寄与率は十分大きく、高い1次元性が確認された。また、項目特性曲線に基づいてモデルの適合性を検討したところ、2-パラメタロジスティックモデルはデータによく適合していることが示唆された。テスト情報関数を求めたところ、中程度の能力レベルを中心に広い範囲の能力レベルで十分なテスト情報量が得られることが分かった。

4. 信頼性の検討：再テスト信頼性を検討するため、1996年11月に、本テストを実施した905名のうち301名の高校生を対象に、NRTを実施した。再テスト法による信頼性係数の推定値は0.81であり、十分高い信頼性が確認された。

また、テスト情報関数に基づいて測定の標準誤差の検討を行ったところ、同様に高い信頼性が確認された。

5. 妥当性の検討：構成概念妥当性を検討するため、1996年11月に、再テストを実施していない604名のうち362名の高校生を対象に、Otis-Lennon School Ability Test (OLSAT 7) を実施した。OLSAT 7の数的推理能力を測定する下位検査とNRTの得点間の相関係数は0.66であり、NRTの数的推理能力テストとしての妥当性が確認された。

【語彙理解能力テストの開発】

1. 項目作成：ミャンマー国語辞典、高等学校で用いられる国語の教科書、高校生向けの学習雑誌などを参考に、語彙理解能力を測定するための多枝選択式の項目を245項目作成した。そして、それらについて6名のミャンマー語の研究者や専門家の検討を受け、さらに、5名の高校生にすべての項目を実施し、全体の回答時間、各項目に対する回答速度、テスト後の感想などの情報を収集した。これらの情報をもとに、不適切と判断された語彙などいくつかの語彙を修正して、245項目からなる予備テストを作成した。

2. 予備テスト：1998年7～8月に、240名の高校生を対象に予備テストを実施した。得られたデータをもとに項目正答率などを検討し、また、因子分析による1次元性の確認、IRTモデルの適用による項目パラメタの推定などを行い、不適切と判断された項目を削除して、139項目からなる本テストを作成した。

3. 本テスト：1998年8～9月に、多段階比例抽出法で抽出した1,025名の高校生標本を対象に本テストを実施した。実施に際しては、あて推量による誤差を極力なくすため、正答が分からない場合は無回答とすること、このテストの結果は学校の成績には無関係であることを十分に学生に説明した。得られたデータをもとに因子分析による1次元性の確認、項目パラメタの推定を行い、因子負荷が0.30以下の項目、項目パラメタの推定値が他の項目と重複する冗長な項目を削除し、最終的に52項目からなる語彙理解能力テスト (Myanmar Vocabulary Comprehension Ability Test; MVCT) を作成した。因子分析の結果をみると第1因子の寄与率はきわめて大きく、非常に高い1次元性が確認された。また、項目特性曲線に基づいてモデルの適合性を検討したところ、2-パラメタ・ロジスティックモデルはデータによく適合していることが示唆された。テスト情報関数を求めたところ、中程度の能力レベルを中心に広い範囲の能力レベルで十分なテスト情報量が得られることが分かった。

4. 信頼性の検討：再テスト信頼性を検討するため、1999年1月に、本テスト

を実施した 1,025 名のうち 873 名の高校生を対象に、MVCT を実施した。再テスト法による信頼性係数の推定値は 0.75 であり、高い信頼性が確認された。また、テスト情報関数に基づいて測定の標準誤差の検討を行ったところ、同様に高い信頼性が確認された。

5. 妥当性の検討：内容妥当性を検討するため、1999 年 10 月に、ミャンマー語の専門家および高等学校の国語（ミャンマー語）教師 24 名に、MVCT に関するアンケート調査を実施した。内容は、日常生活におけるその語彙の重要性、高等学校レベルの語彙としての妥当性、選択枝の適切性、内容領域の適切性の 4 つの観点から各項目を評定し、さらに、テスト全体についても評価してもらうものである。結果は概ね高い評価であり、MVCT の語彙理解能力テストとしての妥当性が確認された。ただし、アンケートの中には有益な意見もいくつか述べられており、今後さらにテストを改良していく必要性も示唆された。

【数的推理能力および語彙理解能力の教科成績との関係についての検討】

基準となる教科成績が、学校ごとに異なるテストで測定されていることへの対応として、次のような方法を考えた。なお、教科テストとして不適切と判断された月例教科テストの成績は事前に除外した。まず、学校ごとに被験者を適当に 2 群に分け、各学校の一方の群を用いて、ベイズ統計学的手法である π 集団回帰を行い、学校に対応する回帰係数の推定値を求めた。そして、その回帰予測式を用いて、もう一方の群における予測値を求め、もう一方の群の各学校のテスト得点の予測値と実際の教科成績との交差妥当性指標の値を推定した。以上を数十回繰り返して、学校ごとに、交差妥当性指標の推定値の中央値を求め、さらに、その中央値を計算し全体の交差妥当性指標の推定値とした。

この方法を用いて、数的推理能力と数学および理科の成績との交差妥当性指標、また、語彙理解能力と英語および国語（ミャンマー語）の成績との交差妥当性指標を推定したところ、およそ 0.3~0.5 の範囲の値であった。テストに混入するさまざまな誤差の影響を考えると、この値は満足のいく値であると判断された。

【まとめ】

本研究では、ミャンマーの高校生のための数的推理能力テストおよび語彙理解能力テストの開発を試みた。その結果、十分な信頼性と適度な妥当性をもつ 2 つの能力テストが作成された。今後、実際の教育の場において、これらのテストが利用され、ミャンマーの高校生の知的能力が適切に評価されることが期待される。