

審査の結果の要旨

氏名 山口 一大

教育測定におけるテスト理論モデルとしてはこれまで、対象となるテストで測定される1次元の能力を想定して、個人の能力水準を量的に推定する項目反応モデルが主であった。しかし、そのような、言わば序列づけのためのモデルからは、各個人がどのような理解状況にあり、どのような学習を必要としているかという診断情報を得ることはできない。これに対し、認知診断モデルと呼ばれる一連のモデルは、テストの各項目への正答にどの能力（アトリビュート）が必要かを示すQ行列をもとに、解答データから各個人のアトリビュート習得パターンを推定することによって、そうした診断情報を得ることを目的とする。本研究は認知診断モデルに焦点を当て、その応用に資するべく、理論的・経験的検討を行ったものである。

第1章の序論で認知診断モデルに関する研究をレビューした後、第2章では、近年数多く提案されてきたモデルについて、理論的に整理を行い、新たな観点からモデルの分類を行った。従来の補償型（どれかのアトリビュートがあれば正答できる）・非補償型（すべてのアトリビュートが正答に必要）という分類に対し、モデルのパラメタ数に注目して、簡約モデル、主効果モデル、飽和モデルに分類するという提案である。この分類は、論文後半の実データへのモデル適合の比較において、有効性を発揮する。

上述のように認知診断モデルではQ行列の設定が重要であり、誤設定があるとテストから得られる診断情報が歪められることになる。論文の第3章では、Q行列の誤設定の影響を、様々な要因を操作したシミュレーション実験によって検討している。複数のアトリビュートの間に、あるアトリビュートの習得が他のアトリビュートの習得の前提となるといった階層構造を考慮した点がこれまでの誤設定研究にない独自の特徴である。そして、そこから、直線型、分岐型、収束型という異なる階層構造のうち、特に分岐型において誤設定の影響が大きいという新たな知見を得ている。

第4章と第5章では、種々提案されてきた認知診断モデルを実データに適用し、どのような性質のモデルがデータに最も適合するのかを比較検討している。第4章ではTIMSS2007の算数の日本人データを用い、第5章では範囲を広げて、TIMSSの公式到達度スコアの高い国から低い国まで、多様な解答データを用いてモデル比較を行った。その結果、日本人データでもその他の国のデータでも同様に、パラメタ数が中程度の主効果モデルの適合が、パラメタ数の少ない簡約モデルやパラメタ数の多い飽和モデルよりも、また1次元の能力を想定する従来の項目反応モデルに比べても、良好であることを見出した。第5章ではさらに、認知診断モデルを適用した結果を国ごとに詳細に分析し、どのアトリビュートの習得割合が高いかが国によって異なる等の興味深い知見を得ており、認知診断モデルの有用性を示している。

本研究は理論的検討・経験的検討のいずれも探索的な性質のものであったが、それぞれから認知診断モデルの今後の応用に資する有用な知見が得られており、意義のある学術的貢献となっている。よって、本論文は博士（教育学）の学位を授与するにふさわしい水準にあるものと判断された。