

大学ガバナンスの効果測定に関する探索的研究

—項目反応理論を用いたガバナンス変数の妥当性の検証—

中 尾 走

大学ガバナンスの効果測定に関する探索的研究

—項目反応理論を用いたガバナンス変数の妥当性の検証—

中尾 走*

1. はじめに

本稿の目的は、項目反応理論 (Item Response Theory, 以下 IRT) を応用して大学のガバナンスの測定の妥当性を再検討することにある。本稿でこのような目的を設定したのは以下の二つの理由による。

第一の理由としては、昨今の一連の大学改革の中でガバナンス改革が大きく取り上げられている点である。そもそも大学のガバナンス自体を取り上げた研究はほとんどなく (黒羽, 1998; 羽田, 2005), ガバナンスを含む大学組織研究自体, 十分な蓄積がないとされる分野である (村澤, 2009)。高等教育研究に比して, 政策レベルでは古くからガバナンスや管理運営の問題は取り上げられてきた。学生紛争以降, 全学的な学園秩序維持という名目で1963年中央教育審議会答申 (通称, 38答申) では, 大学自治を教授会自治から大学全体の自治へと確立しようとした (羽田, 2014)。その中で, 学長は全学的な責任者と位置づけられ, 学内のアクターの職務・権限の明確化が提言されている (羽田・金井, 2010)。その後, 半世紀が経ち, 学長リーダーシップ論は現在文部科学省関係の審議会だけでなく, 教育再生実行会議の第三次提言においても大学ガバナンスの機能強化が取り上げられるほどになった (教育再生実行会議, 2013)。その後, 現在の学長リーダーシップ論へと繋がるのは2013年の中教審大学分科会組織運営部会で学長が最終的な意思決定者と位置づけられたことによる (中教審大学分科会組織運営部会, 2013)。このように, 政策レベルにおいて学長の権限はますます強くなってきたわけだが, 研究ではその実態分析が蓄積されてきた (広島大学高等教育開発センター編, 2007; 国立大学財務・経営センター, 2007, 2010など)。これらの実態分析では, 学長を含む執行部への権限集中は広く認知されており, 特に国立大学でその傾向は顕著である事, 組織運営の在り方など, 組織の階層レベルで相違があることなどが明らかにされてきた。近年は, このような実態把握から一步進んで, 学長のリーダーシップを含む権限を一箇所に集中させることの効果を取り上げることへの着目がされるようになった (村澤, 2017; 村澤・中尾・松宮等, 2019)。例えば, 村澤 (2017) では権限の集中度をガバナンスとして教育成果との関係について, 村澤・中尾・松宮 (2019) では研究成果との関係を検討している。これらの実証分析では, いずれもガバナンス≒権限を集中させたことの効果は確認されていないが, そもそもガバナンス自体が適切に変数化できていない可能性を残している。なぜなら, 質問紙におけるガバナンスの測定項目を全て等価なものとして総合得点化し, ガバナンス≒権限の集中度としているが, 権限の集中度が低い大学と

* 広島大学大学院教育学研究科教育学習科学専攻 (高等教育学)

高い大学で権限の集中度の内実が異なる可能性があるからである。例えば、人事・資金配分・学部・学科の改組の合計3点と進路指導・授業時間の決定・カリキュラムの策定の合計3点を同じ権限の集中度と捉えていいものかという疑問が残る¹⁾。この問題に対し、IRTを用いることで対処可能ではないかと考える。そこで、本稿ではIRTの応用により、大学のガバナンスの測定に関する再検討を行う。本項でガバナンスを取り上げる理由は、ガバナンスの関連領域である社会学的権力理論は、権力概念の定義、本質規定、測定法やそれに関連する操作的定義が権力理論の第一歩であったという(志田, 2000)。高等教育におけるガバナンス研究に関しても、その第一歩として、測定法や関連する操作的定義を固めていくことが必要であるからである。

第二の理由は、近年、政策レベルで根拠に基づく政策立案(Evidence-based policy making, 以下EBPM)が求められており、その中で測定法の重要性を問題提起することにある。周知の通り、政策立案一般に対して「科学的根拠に基づくこと」が要請されており、高等教育政策に与える影響も無視できない。特に、高等教育という対象は政策からの影響を受けやすく(吉田, 2013)、この影響を直接受ける可能性がある。EBPMブームは、これまで何度もあったとする見方もあるが(白川, 2019)、昨今のEBPMは“狭義のEBPM”と呼ばれ(加納・林, 2019)、その特徴は、エビデンスの質に序列がつけられている点にある。“狭義のEBPM”の重要な役割の一つとして取り上げられるのが、客観的手法に基づく政策の事後検証である²⁾。しかし、教育の分野では例えば大学進学の効果など教育の効果は一様ではなく(中澤, 2017)、ほとんどの指標が厳密には測定しきれない(中村, 2018)。これまでの大学改革には客観的に測定し難いものも多く含まれ、今回分析で取り上げるガバナンスもその一つである。このとき、“狭義のEBPM”が想定するような政策の事後検証を行うには指標の適切な構成概念化は重要課題である。しかし、“狭義のEBPM”では分析手法ばかりが前面に出され、政策の事後検証に用いる構成概念が適切に測定されているのかという議論は皆無に等しい³⁾。また、高等教育研究はその多くが量的研究であるにも関わらず(立石・丸山・速水・松宮・中尾・村澤, 2019)、IRTなどの測定論を用いた指標の妥当性を検討したものは木村(2012)を除き、数少ない⁴⁾。本稿で用いるガバナンス以外にも学生調査、大学教員調査などで海外との比較研究のために用いた諸指標についても、日本の現状を測定するものとして適切かどうかを分析・検討した研究は数少なく、この点については反省的になるべきであろう。

そこで、本稿では主に心理学領域において研究の蓄積があるIRTを応用して、ガバナンスを測定してきた項目の妥当性を検討する。続く第2節では、用いるデータと変数の特徴、分析方法の詳細を解説し、合わせてIRTを本項で用いる妥当性を説明する。第3節では、実際の分析結果と考察を、最後に第4節ではまとめと今後の課題、含意を提示する。

2. データと方法

(1) データと変数

本稿では、「大学における教学ガバナンスとその効果に関する調査研究」の二次データ(以下、ガバナンスデータ)を用いた。調査期間は2015年1月から5月であり、全国の大学の学部長2,216人

に対して質問紙による郵送調査を行った。回収数は1,044人、回収率は47.1%であった。本稿で着目する変数であるガバナンスは、大学経営全般に関することの18領域（例えば、教員の人事や副学長の任免など18項目）と教学に関することの14領域（例えば、全学の入試選抜方法や教員の担当授業の決定など14項目）の32の意思決定が必要とされる領域を設定し、その領域の意思決定に対する13の主体の影響力の強さを学部長の視点から測定している⁵⁾。これらの回答結果は、32領域×13主体のマトリクスになっており、32領域それぞれにおける最も影響力の大きい主体を答えてもらっている。ゆえに、IRTで使いやすいような、ある領域の意思決定をどの主体がしているかというマトリクスのデータフレーム構造となっている。表1には、13の主体ごとに32の領域をそれぞれ1点とした0~32点のガバナンス変数の基礎統計量を示している。また、分析に際して、ガバナンス項目に回答しているサンプルのみを分析対象とした⁶⁾。

表1 ガバナンス変数の基礎統計量

| 主体 | n | Mean | SD | min | max |
|---------------|------|-------|-------|-----|-----|
| 政府 | 1045 | 0.096 | 0.369 | 0 | 3 |
| 地方自治体 | 1045 | 0.201 | 0.909 | 0 | 11 |
| 産業界 | 1045 | 0.014 | 0.188 | 0 | 4 |
| 学長 | 1045 | 6.413 | 5.143 | 0 | 30 |
| 理事長 | 1045 | 2.489 | 3.731 | 0 | 25 |
| 理事・副学長 | 1045 | 1.421 | 2.893 | 0 | 21 |
| 理事会・役員会等 | 1045 | 3.237 | 3.38 | 0 | 28 |
| 教育研究評議会・評議員会等 | 1045 | 1.484 | 2.547 | 0 | 15 |
| 部局長 | 1045 | 3.592 | 3.996 | 0 | 23 |
| 教授会 | 1045 | 7.313 | 5.531 | 0 | 23 |
| 教員 | 1045 | 3.896 | 3.468 | 0 | 24 |
| 事務組織 | 1045 | 2.145 | 2.465 | 0 | 21 |
| 同窓会組織等 | 1045 | 0.012 | 0.123 | 0 | 2 |

表1より、多くの大学が権限を学長に集め、影響力が大きいことが分かる。学長への権限の集中はこれまでの先行研究とも類似の結果である（根岸，2015）。しかし、平均だけで比較すると教授会が最も高く、学長へ権限が集まっているにも関わらず、未だに教授会の影響力の大きさは無視できないと言える。実際の分析には、近年の大学改革においては学長のリーダーシップと教授会の役割の明確化が何度も取り上げられ（中央教育審議会大学分科会組織運営部会，2013）、政策上の関心の高さから学長と教授会の影響力の強さを分析対象とした。

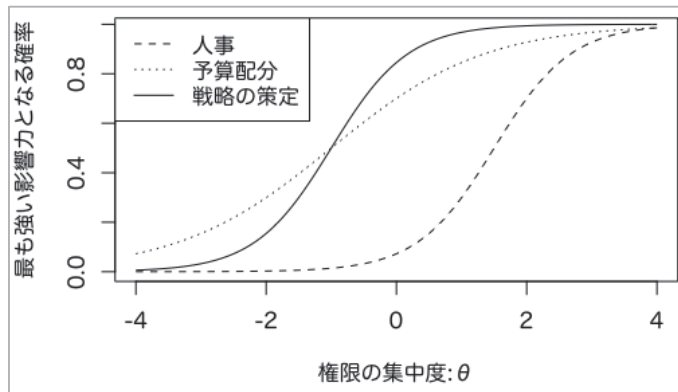
(2) IRT のガバナンスへの応用

本稿では、ガバナンス変数の再構成を行うにあたって、IRTを用いる。テストを事例に学力を測定する文脈で説明するとIRTは、項目識別力（Item discrimination, 以下識別力）と項目困難度（Item difficulty, 以下困難度）という二つのパラメータによってそのテストで適切に学力が測定できてい

るのかの妥当性を問う。識別力とは、そのテストの項目の正解／不正解で学力の高い／低いを識別できるかを示すものである。困難度とは、そのテストの項目に対する正解のしにくさである。この二つのパラメータの推定する際、古典的テスト理論（Classical Test Theory, 以下 CTT）では標本依存性（sample dependence）と項目依存性（item dependence）という二つの問題点を抱えていた。標本依存性とは CTT で推定される困難度や識別力というパラメータがそのテストを受験する集団に依存して決定されるということである。項目依存性とは、CTT で推定されるパラメータがそのテストの項目自体に依存して決定するということである。これらの CTT の二つの問題点を克服するものとして、IRT は誕生した経緯をもつ。IRT の大きな特徴は、困難度と受験者の能力を別々に推定することで分離して表現できるところにある。そのため、CTT で大きな問題点とされてきた、標本依存性と項目依存性が解決可能となる⁷⁾。つまり、IRT を用いることで、項目によらない測定（item-free measurement）と受験者によらない測定（sample-free measurement）が可能となる。この性質は、本稿で IRT を用いる上で相応しい性質である。なぜなら、クロンバックの α 係数や因子分析など指標の妥当性や構成概念を変数化するこれまでの分析方法においても標本依存性と項目依存性の問題を抱えており、得られたデータのサンプルに依存して、妥当であったり妥当でなかったりと結果が変化するので、妥当性を検討したと云い難いからである。例えば、葛城（2007, 2010）など妥当性を検討したと主張する論文はあるが、標本依存性や項目依存性の問題点を克服できていない。

IRT に関するこれ以上の詳細な説明は専門書に譲り、大学のガバナンスへ応用するに当たって、図1に解説のための曲線を図示した。図1は3つの領域の項目特性曲線（Item Characteristic Curve, 以下 ICC）の事例である。それぞれの項目のパラメータは、人事が $a=1.0$, $b=1.5$, 予算配分が $a=0.5$, $b=-1.0$, 戦略の策定が $a=1.0$, $b=-1.0$ となっている。この a と b がそれぞれ識別力、困難度のパラメータであり、それぞれ曲線の傾きと曲線の位置を示している。図1の例で言えば、予算配分と戦略の策定では b の値は同じであるが、 a の値が異なる。これは予算配分と戦略の策定では傾きが異なることを示す。人事と戦略の策定は $a=1.0$ と同じであるため、曲線の位置は異なっている（ $b=-1.5$ と $b=-1.0$ で b の値が異なる）が傾きは同じであることが読み取れる。横軸のパラメータ（ θ ）は一般的に、 -4 から $+4$ までの範囲をとり、ある主体に権限が集まれば集まるほど、人事・予算配分・戦略の策定の決定に対して最も影響力の強い主体となる確率が高くなるとする。また、 0 は平均的な権限の強さを示す。

人事と戦略の策定を比較すると、権限がより集中しなければ、決定に影響を及ぼさないのは人事の方である。逆に、権限が集中しなくとも、戦略の策定に影響を及ぼす確率は高くなることが分かる。また、戦略の策定と予算配分を比較すると、権限の集中度によってその影響力の大きさに逆転が起きている。権限が集中していないとしても、予算配分に対して最も影響力が大きくなる可能性が高く、戦略の策定は権限がある程度集中していくことで、最も影響力が強くなる。実際のデータからこのような逆転現象が生じていれば、権限が弱い主体に対してと強い主体に対して質問項目を変更した方がよいことを示す。また、推定パラメータである識別力が極端に低い領域の質問項目があれば、その質問項目は権限の集中度を測定するのに必要ではない質問項目であることを示す。このように ICC には、横軸のパラメータ（ θ ）ごとの確率と識別力、困難度の情報が付随している。



(加藤・山田・川端, 2014) 参考に筆者作成

図1 大学のガバナンスへの応用例

3. 分析と考察

(1) IRT 適用の前提

本稿で、IRT を用いて分析するにあたって、二つの仮定を満たすことが望ましい。それは、一次元性 (unidimensionality) と局所独立性 (local-independence) である。一次元性とは、32の領域の質問項目で一つのガバナンスという構成概念を測定していることを示す。もし、この仮定が満たされないのであれば、32の質問項目で複数のガバナンスを測定していることを示し、多次元 IRT モデル (Multidimensional IRT model, Reckase, 2009) などを適用することが望ましい。局所独立性に関しては、ガバナンスという変数を構成する際に、32の領域の質問項目が独立であるかというものである。テストの分析に用いられる場合は、問1が解けなければ問2が解けない場合など、この仮定を満たさない可能性がある。局所独立性が満たされない場合は、推定パラメータの標準誤差が大きくなることが知られている (登藤, 2010)。これは IRT を用いる上での重要な仮定であるが、村澤 (2017) や村澤・中尾・松宮 (2019) でガバナンス変数を構成する上で暗黙のうちに前提とされていた仮定でもある。これらの仮定が満たされているかどうかの確認が行われていないため、検証すべき事項である。

加藤・山田・川端 (2014) では、一次元性の確認は因子分析を応用した方法を、局所独立性の確認は Yen (1993) に従い Q3統計量を事例として用いている。局所独立性の仮定に関しては、大きく分けて三つの事例が考えられるが、本稿の場合、項目連鎖 (item chain) や文脈依存 (context dependence) など項目前後の回答から直接的、間接的影響を受けることは想定しづらい。しかし、測定対象のガバナンス θ と異なる概念である θ' の影響を受けることによって、局所独立性の仮定が満たされていない可能性がある。

一次元性の確認は、因子数ごとに固有値を計算し、様々な基準で因子数を決定し、1因子であれば一次元性の仮定が満たされているとする。詳細な分析結果は紙幅の都合上、省略するがスクリー

基準, ガットマン基準, 平行分析 (parallel analysis), MAP (Minimum Average Partial) 基準の四つの方法で検討を行うと, 学長のガバナンスと教授会のガバナンスはどちらも1因子から5因子まで想定される結果となった。最終的には分析者が解釈可能性の観点から決定を行うことが推奨されているが (堀, 2005), 本稿で用いたガバナンス変数は一次元性が満たされていると言えなくはないが, 強くは言い切れないということである。

次に, 局所独立性の仮定について, Q_3 統計量から検討を行う。紙幅の都合上, Q_3 統計量の結果も省略する。 Q_3 統計量は評価の目安として絶対値が0.2以上であれば局所独立性が満たされていないことを示すという一つの基準を用いられるが (加藤・山田・川端, 2014), 学長と教授会のガバナンスについて, 完全に局所独立性の仮定を満たされているとは言い難い結果であった。本節での分析結果から, 一次元性の仮定については, 満たされているとも満たされていないとも両方の解釈が可能である。また, 局所独立性の仮定は, Q_3 統計量に基づいて判断すると満たされているとは言い難い。そのため, IRT を用いると推定パラメータの標準誤差が大きくなる可能性があるが, 本稿では先行研究で用いられてきたガバナンス変数の検証という意味合いが大きいいため, 次節以降そのまま IRT を適用することとする。

(2) IRT による分析

本節では, IRT を用いて質問項目が大学のガバナンスを適切に測定できているかを明らかにする。表2, 3はその推定されたパラメータであり, 図2から図5はそれらの ICC である。図2から図5の全てのグラフは横軸に権限の集中度, 縦軸に最も影響力が強い主体となる確率を示している。また, 図2から図5の判例はそれぞれ表2, 3の記号を示している。

表2 IRTで推定されたパラメータ

| | 大学経営全般 | 学長 | | 教授会 | |
|-----|---------------------------|---------|--------|--------|-------|
| | | 困難度 | 識別力 | 困難度 | 識別力 |
| 1_a | 全学の戦略, 目標計画等 | -0.055 | 0.561 | 3.024 | 1.753 |
| 1_b | 学部の戦略, 目標計画等 | 1.452 | 1.116 | 0.884 | 1.282 |
| 1_c | 競争的資金の公募プロジェクト・プログラムの企画立案 | 1.471 | 1.137 | 2.919 | 1.358 |
| 1_d | 教員採用枠・数 | 1.273 | 0.875 | 2.260 | 1.200 |
| 1_e | 個々の教員人事 | 1.239 | 1.655 | 0.150 | 1.276 |
| 1_f | 教員の業績評価 | 1.197 | 1.700 | 1.203 | 1.200 |
| 1_g | 学長の任免 | 6.916 | 0.643 | 6.202 | 0.500 |
| 1_h | 理事長の任免 | -20.659 | -0.131 | 20.192 | 0.277 |
| 1_i | 副学長の任免 | -1.849 | 0.148 | 9.152 | 0.481 |
| 1_j | 部局長の任免 | 0.837 | 0.788 | 1.839 | 0.669 |
| 1_k | 学科・専攻の長の任免 | 1.080 | 1.324 | 1.082 | 1.189 |
| 1_l | 学部・研究科の改組・再編 | 0.911 | 1.428 | 1.009 | 1.287 |
| 1_m | 全学の評価方針 | 0.129 | 1.459 | 2.659 | 1.373 |
| 1_n | 学部の評価方針 | 0.965 | 1.790 | 1.059 | 1.442 |
| 1_o | 全学予算の配分 | 1.281 | 0.868 | 4.587 | 1.060 |
| 1_p | 学部予算の配分 | 1.580 | 1.452 | 1.740 | 0.958 |
| 1_q | 全学の事務の組織構成・役員人事 | 2.204 | 0.865 | 8.888 | 0.598 |
| 1_r | 部局の事務の組織構成・役員人事 | 2.177 | 1.104 | 5.063 | 0.827 |

表3 IRTで推定されたパラメータ

| | 教学 | 学長 | | 教授会 | |
|-----|-------------------------|-------|-------|--------|-------|
| | | 困難度 | 識別力 | 困難度 | 識別力 |
| 2_a | 全学の入試試験方法・合否の方針 | 0.578 | 2.357 | 1.425 | 1.047 |
| 2_b | 学部の入試試験方法・合否の方針 | 1.146 | 2.534 | -0.092 | 2.033 |
| 2_c | 全学の教育方針・カリキュラム・プログラムの策定 | 0.567 | 2.145 | 1.305 | 1.324 |
| 2_d | 学部の教育方針・カリキュラム・プログラムの策定 | 1.517 | 2.506 | -0.209 | 2.841 |
| 2_e | 教員の担当授業の決定 | 1.919 | 3.238 | -0.035 | 2.811 |
| 2_f | 教員の授業時間の決定 | 1.856 | 2.879 | 0.084 | 2.439 |
| 2_g | 成績評価の基本方針 | 1.478 | 2.763 | 0.283 | 2.202 |
| 2_h | 授業評価の基本方針 | 1.357 | 2.218 | 0.358 | 2.225 |
| 2_i | 進路指導 | 2.253 | 2.849 | 1.124 | 2.216 |
| 2_j | 学生生活(福利厚生)支援 | 2.140 | 2.732 | 1.242 | 2.214 |
| 2_k | 学習支援 | 2.050 | 4.140 | 1.078 | 2.202 |
| 2_l | FD | 1.614 | 2.054 | 0.642 | 1.609 |
| 2_m | 教学に関する情報・データ収集と分析(IR) | 1.764 | 1.803 | 1.340 | 1.610 |
| 2_n | 授業に関わる学習時間(予習復習)の策定・想定 | 2.015 | 2.520 | 0.939 | 2.040 |

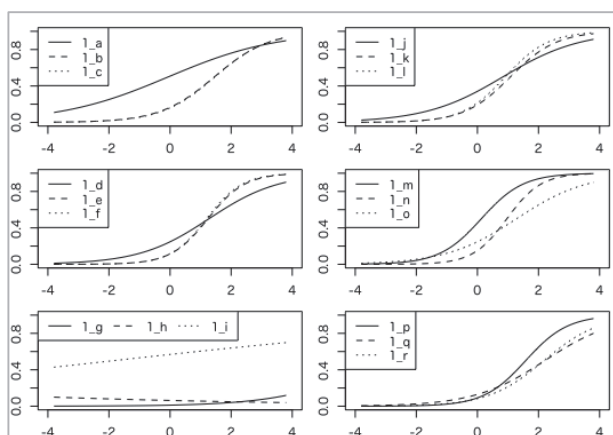


図2 学長の権限の集中度と大学経営全般に関する18の質問項目のICC

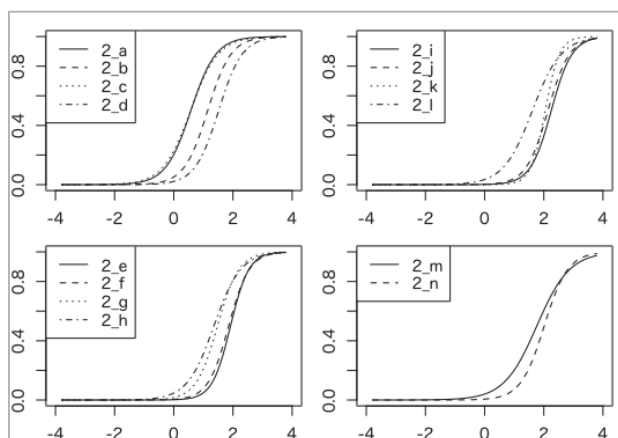


図3 学長の権限の集中度と教学に関する14の質問項目のICC

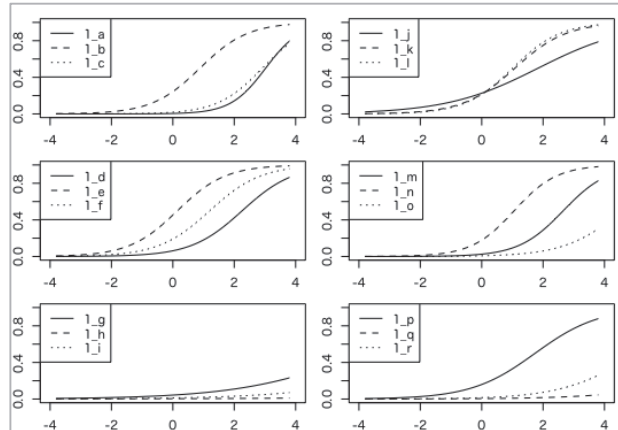


図4 教授会の権限の集中度と大学経営全般に関する18の質問項目のICC

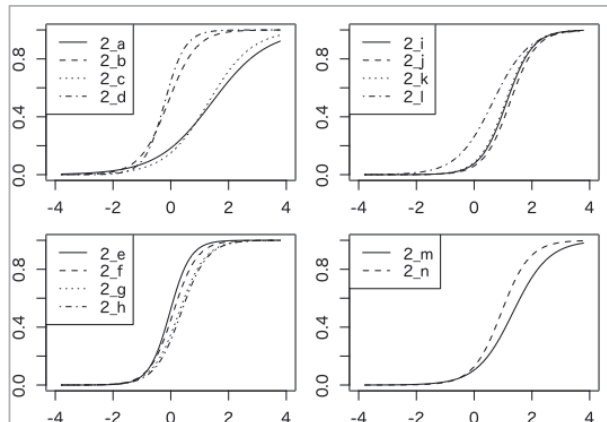


図5 教授会の権限の集中度と教学に関する14の質問項目のICC

図2, 3は学長の権限の集中度に関する ICCである。大学全般に関する領域の質問項目では g, h, i の3つの領域の質問項目が他の領域の質問項目と大きく様相が異なる。これらはそれぞれ、学長の任免、理事長の任免、副学長の任免である。学長と理事の任免についてはどれだけ学長に権限が集まったとしても最も影響力が強い主体となる確率が極端に低く、逆に副学長に任免については、学長に権限が集中する／しないに関わらず学長が最も影響力が強い主体である確率が非常に高い。これら3つの領域に関する質問項目は、学長の権限の集中度を識別するには適切な質問項目ではなく、変数の再構成をする上で他の質問項目と等価として扱うべきではないことが分かる。加えて、j, k, lに着目すると横軸の0~1付近で曲線が逆転していることがわかる。つまり、これらの項目は権限の集中度の強弱によって最も影響力が強い主体となる確率が逆転するということである。一部のアクターに、権限を集中させればさせるほど、より多くの領域の意思決定者になるわけではなく、一

部のアクターに集中させすぎないように、選択的に権限を集中させていることを示す。これは、権限の強さだけを分析対象にするのではなく、ガバナンス変数はどの領域を決定するかも重要であるということである。また、図3の教学に関する14の領域では、図2と異なり全てが似たような曲線の形状をとっている。アルファベット順に a から n に向かうにつれて、曲線が徐々に右に寄っていくことが分かる。これは、a から n の領域に向かうに従って、学長が最も影響力が強い主体となるためには、学長により多くの権限を集めなければならないことを示している。次に、図4、5は教授会の権限の集中度に関する ICC である。教授会に関しては、どれだけ権限が集中しても、学長の任免、理事長の任免、副学長の任免について最も影響力のある主体となることはほとんどない。また、全学予算の配分、全学の事務の組織構成・役職人事、部局の事務の組織構成・役職人事についても類似の曲線となっている。教授会に関しては、上記の6つの項目で最も影響力のある主体になる確率は0に近く、教授会のガバナンスのみを測定するのであれば、質問項目として必要ないことを示す。表1では教授会のガバナンスの平均が7.313で、13の主体の中で最も高い。その中身は、6項目でほとんどの確率で最も影響力のある主体にならないので、実質は32から6を引いた26項目でガバナンスを測定されていると認識しても間違いはない。また、学長は最も影響力のある主体になる確率がほとんど0である項目が2つあり、30項目でガバナンスが測定されている⁸⁾。

4. まとめと今後の課題

本稿では、IRT を応用して先行研究で用いられてきたガバナンス変数の再検討を行った。得られた主な知見は以下の三つである。

第一に、一次元性と局所独立性が未充足である点である。先行研究では32の測定項目を全て等価なものとして総合得点化することでガバナンス変数として用いてきたが、暗黙の内に一次元性と局所独立性の過程が満たされているという前提の変数構成をしていた。しかし、実際には32項目で測定されてきたものが、いくつかの下位概念から成り立っている可能性がある。この時、多次元 IRT モデル (Multidimensional IRT model, 星野, 2001 ; Reckase, 2009) などを用いての分析は今後の課題である。

第二に、32項目を全て等価なものとして扱えないという点である。図2から図5で示されている通り、一つの主体に着目すると、識別力 (曲線の傾きが小さすぎる／大きすぎる) が低い項目がいくつかあり、ガバナンスの測定項目として妥当ではないことを示している。もちろん、この調査票では本稿で着目した学長と教授会以外の主体についても尋ねているので、この結果から質問項目から外すべきとの結論は導くことは出来ない。しかし、本稿での分析結果に即せば、権限の集中度≒ガバナンスとして変数に用いるには、等化 (equating) などを行ったうえで、学長と教授会のガバナンス変数を作成すべきである。本稿では紙幅の都合上、ガバナンス変数を再構成するまでに至らなかったが、IRT の分析結果に基づいてガバナンス変数を再構成し、ガバナンス≒権限の集中度の効果について実証分析することを今後の課題である。

第三に、権限の一元的な集中が、影響力の強さを表すものではないという点である。図2から図5

でいくつかの曲線で逆転する項目がみられる。これは、特定のアクターに権限を集中させたとしても、最も影響力の強い主体となる確率が高くなるとは限らないということである。実際には、権限を集中させれば急激に確率が高まる領域と逆に確率が下がる領域があるため、逆転項目が存在した。つまり、日本の大学ではガバナンス自体が組織デザインであり、政府主導で学長に権限の集中化が進められてきたが（羽田，2019）、実際には学長に権限が集中し過ぎないようにしている可能性がある。組織特性や状況に応じたガバナンスを選択しているという村澤（2017）の知見とも整合的である。政策的関心から、権限の集中度の効果が検討されてきたが、どのようなガバナンスが成果を高めるかを分析することは今後の課題である。

教育の分野では様々な指標を測定すること自体が困難であるため、その変数化に対して様々なアプローチから検討する取り組みが必要であるが、本稿はそのための一つの可能性としてIRTを用いたに過ぎない。「可視化」「数値化」が高等教育の大きなブームになりつつある中で、最も大きな影響力をもつ大学ランキングの指標の妥当性など稿を改めて検討したい。

【付記】

本稿は、JSPS 科研費 JP25301046（基盤研究（B），研究代表者 秦由美子）で収集したデータの提供を受け、使用することをご快諾いただいた。

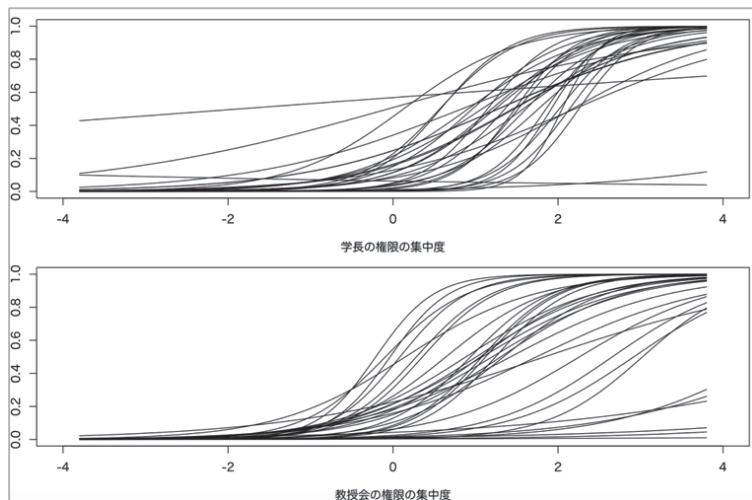
また、本研究推進に際し、以下の資金提供を受けた。JSPS 科研費 JP16H03780（基盤研究（B），研究代表者 村澤昌崇）・JP18K18651（挑戦的研究（萌芽），研究代表者 村澤昌崇）・JP19H00621（基盤研究（A），研究代表者 小林信一）。ここに記して感謝申し上げる次第である。

【注】

- 1) 村澤（2017）、村澤・中尾・松宮（2019）にもあるように、権限の行使が法・規程どおりには発動しえない点を踏まえ、権限の所在やその軽重（どの権限かどこに行使され、どのような結果をもたらすのか）を正確に測るよりも、権限が人々にどのように意識・認識されているかを測ることに主眼があるのであれば、測定項目の単純な加算による権限の包括的把握には意義はある。
- 2) これは、“狭義のEBPM”が一部の経済学者によって牽引されているにも関わらず（小林，2018）、その手法が誘導型推定に偏っていることから自明である。
- 3) 呉・島・西村（2019）は“狭義のEBPM”で推奨される手法に引っ張られたことで、連続量のもを2値変数として用いた典型である。連続量のもを2値として用いることは検定力が下がり（Reinhart 2015=2017）、交絡因子を適切に区別できないことが知られている（Maxwell & Delaney, 1993）。交絡を統制する目的で傾向スコアを用いながらも、結果的に交絡を統制できない分析枠組みとなってしまった良い例である。
- 4) 昨今の大学入試改革の議論も踏まえて、テストに関するIRTの限界についての言及（荒井，

2015) や IRT を用いた学力問題に関する論考 (川口, 2015) など遅ればせながらいくつか散見されるが、心理学の研究蓄積に比して、その他の社会科学の分野において社会調査一般で、測定指標の妥当性の検討自体が一般的なものとなっているとは言い難い。

- 5) 質問票については、村澤 (2017) 参照のこと。加えて、学部長の主観による回答をガバナンスとして測定した理由についても、村澤 (2017), 村澤・中尾・松宮 (2019) を参照のこと。
- 6) 村澤 (2017) において、設置者や大学の威信などの主な属性別にみると、回収結果が母集団分布から大きく偏っていないことは確認済みである。
- 7) IRT を用いれば、必ず標本依存性と項目依存性の問題点が解決されるわけではない (加藤・山田・川端, 2014)。
- 8) 実際には、「誰がどのような権限を持っているか」という学部長の認識を測るという点に対して重み付けなど細かな違いは必要ない。しかし、等化で用いたガバナンス変数の効果を測定する際には、補図1のような逆転項目がいくつか存在する場合、ガバナンス \approx 権限の集中度の効果を適切に推定できない可能性がある。なぜなら、被説明変数 y の変化量は説明変数 x が0から1と10から11で1変化したときの変化量と異なるからである。ただ、これらはガバナンスの認識論の違いでもあるため、先行研究に大きな問題があるわけではない。また、補図1は図2と3、図4と5の学長と教授会の32の領域をそれぞれ傾きの逆転が分かりやすいように一つの図にまとめたものである。



補図1 学長と教授会のICC

【参考文献】

- 荒井克弘 (2015) 「ますます不透明な『新テスト』の行方」『学研・進学情報』12, 2-5頁。
 加藤健太郎・山田剛史・川端一光 (2014) 『Rによる項目反応理論』オーム社。

- 加納寛之・林岳彦（2019）「環境分野へのEBPMの導入に向けての概念整理」『第144回関西公共政策研究会発表資料』。
- 川口俊明（2015）「項目反応理論による学力調査の再分析」『福岡教育大学紀要』64（4），1-12頁。
- 木村拓也（2012）「大学満足度の学年変化とその規定要因の探索」『Journal of Quality Education』4，73-91頁。
- 教育実行再生会議（2013）『これからの大学教育等の在り方について（第三次提言）』。
- 葛城浩一（2007）「就職率の教育成果指標としての妥当性」『大学論集』第38集，209-220頁。
- 葛城浩一（2010）「アウトカム指標のあり方を考える」『大学論集』第41集，439-454頁。
- 黒羽亮一（1998）「研究の分野と人を限定することなく」『高等教育研究』1，113-115頁。
- 国立大学財務・経営センター（2007）「国立大学法人の財務・経営の実態に関する全国調査」（2006年実施：国立大学財務・経営センター研究部）。
- 国立大学財務・経営センター（2010）「国立大学法人の経営・財務の実態に関する研究報告書」（2009年実施：国立大学財務・経営センター研究部）。
- 呉書雅・島一則・西村君平（2019）「日本学生支援機構貸与型奨学金の受給が生活時間に与える影響」『高等教育研究』22，207-229頁。
- 小林信一（2018）「科学的根拠に基づく政策」『科学』88（11），1149-1156頁。
- 志田基与師（2000）「権力理論の説明様式」『理論と方法』15（2），299-312頁。
- 白川展之（2019）「根拠に基づく政策におけるエビデンスとは何か：国際的動向と日本の現実」『2018年度広島大学高等教育研究開発センター第8回公開研究会発表資料』。
- 惣脇宏（2010）「英国におけるエビデンスに基づく教育政策の展開」『国立教育政策研究所紀要』139，153-168頁。
- 立石慎治・丸山和昭・速水幹也・松宮慎治・中尾走・村澤昌崇（2019）「高等教育と計量分析」『教育社会学研究』104，29-53頁。
- 中央教育審議会大学分科会組織運営部会（2013）「大学のガバナンス改革の推進について（審議まとめ）」。
- 登藤直弥（2010）「局所独立性の仮定が満たされない場合の潜在特性推定への影響」『日本テスト学会誌』6，17-28頁。
- 豊田秀樹（2002）『項目反応理論—入門編—』朝倉書店。
- 中澤渉（2017）「教育社会学と計量分析—到達点と今後の課題」本田由紀編『教育社会学のフロンティア1 学問としての展開と課題』109-126頁。
- 中村高康（2018）『暴走する能力主義』ちくま新書。
- 根岸佳代（2015）「学長のリーダーシップ観」『IDE 現代の高等教育』567，63-69頁。
- 羽田貴史（2005）「大学管理運営論」有本章他編『高等教育概論—大学の基礎を学ぶ』ミネルヴァ書房，30-40頁。
- 羽田貴史（2014）「教育マネジメントと学長リーダーシップ論」『高等教育研究』17，45-63頁。
- 羽田貴史（2019）『大学の組織とガバナンス』東信堂。

- 羽田貴史・金井徹 (2010) 「国立大学長の選考制度に関する研究—選挙制度の定着と学長像」『日本行政学会年報』36, 158-175頁。
- 広島大学高等教育研究開発センター編 (2007) 『大学の組織変容に関する調査研究』COE 研究シリーズ27, 広島大学高等教育研究開発センター。
- 星野崇宏 (2001) 「多次元項目反応理論での相関のある特性値の線形結合に関するテスト情報関数」『教育心理学研究』49, 491-499頁。
- 堀啓造 (2005) 「因子分析における因子数決定法」『香川大学経済論叢』77, 35-70頁。
- 村澤昌崇 (2009) 「日本の大学組織—構造・機能と変容に関する定量分析」『高等教育研究』12, 7-28頁。
- 村澤昌崇 (2017) 「大学におけるリーダーシップと環境・戦略・組織特性そして成果—全国学部長アンケート調査に依拠して—」広島大学高等教育研究開発センター編『大学運営におけるリーダーシップ—第44回 (2016年度) 研究員集会の記録—』(高等教育研究叢書138), 43-64頁。
- 村澤昌崇・中尾走・松宮慎治 (2019) 「大学の研究生産とガバナンス」『名古屋高等教育研究』名古屋大学高等教育研究センター, 19, 153-169頁。
- 吉田文 (2013) 「グローバリゼーションと大学」広田照幸ほか編『グローバリゼーション, 社会変動と大学』岩波書店, 15-42頁。
- Chen, W. H., Thissen, D. (1997). Local Dependence Indexes for Item Pairs Using Item Response Theory, *Journal of Educational and Behavioral Statistics*, 22, 265-289.
- Maxwell, S. E., Delaney, H. D. (1993). Bivariate Median Splits and Spurious Statistical Significance, *Psychological Bulletin*, 113 (1), 181-190.
- Porter, T. M. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Princeton University press., (=2013, 藤垣裕子訳『数値と客観性—科学と社会における信頼の獲得』みすず書房)
- Reckase, M. D. (2009). *Multidimensional Item Response Theory*, New York, NY: Springer.
- Reinhart, Alex. (2015). *Statistical Done Wrong* No Starch Press., (=2017, 西原史暁訳『ダメな統計学—悲惨なほど完全な手引書』勁草書房)
- Yen, W. M. (1993). Scaling Performance Assessments: Strategies for Managing Local Item Dependence, *Journal of Educational Measurement*, 30 (3), 187-213.

Applicability of Measurement Methods in Higher Education: A Case Study of University Governance

Ran NAKAO*

The purpose of this study is to examine the validity of measurement using Item Response Theory (IRT) in the case of university governance. The main conclusions obtained in this paper are as follows: First, unidimensionality and local-independence were not satisfied. In other words, the research revealed that 32 items consisted of several subordinate concepts. Second, all 32 items cannot be treated as equivalent. Third, the governance of each university means that the concentration of authority is not important and is a combination. From these results, we can conclude that analysis from various viewpoints is necessary to clarify the effect of governance.

* Doctoral Student, Graduate School of Education, Hiroshima University