

大学教育における学生の学修成果の アセスメントのモデル化に向けて

鵜沼秀行*・長谷川 桐**

Towards a Model for Assessing Learning Outcomes of Students in University Education

Hideyuki UNUMA, Hisa HASEGAWA

Abstract

This study modeled causal relationships among the following variables: learning outcomes, the readiness of students for university education, and educational programs. We examined four models explaining longitudinal data of the academic years from 2015 to 2018 for eight educational programs based on the framework of the IEO model (Astin, 1993). The placement test score of students (N=260) before starting the 2015 academic programs and their GPA score just before graduating from the university in 2019 were analyzed using structural equation modeling, by treating the educational programs as dummy variables. The results suggested that students' readiness, as assessed by the placement test scores, and educational programs could explain 13% of the variance in the GPA. We have discussed the validity of the models and proposed suggestions for future modeling studies.

Key Words: University Students, Learning Outcomes, GPA, Placement Test, Structural Equation Modeling

近年、大学教育において学生の学修成果に関する情報や大学としての教育成果についての情報を的確に測定し、さらに教育活動の見直しに活用することは、各方面で広く求められている（中央教育審議会，2018；松下，2017）。特に、私立大学においては「教育の質保証」のために学修成果を可視化するための具体的な手法が議論されてきた（日本私立大学連盟教育研究委員

*教授 認知心理学

**非常勤講師

会, 2019)。

学修成果の評価, 検証, そして教育活動における活用というアセスメントの問題は, 本邦だけでなく, 広く世界的に議論されてきた (OECD, 2008)。米国の大学教育におけるアセスメントの状況を整理した山田 (2005) においては, 履修前基本データ取得用アセスメント, 1年終了時でのアセスメント, 学生の行動・態度・学修スキルについてのアセスメント, などが紹介されている。

このように一般にアセスメントと呼ばれる評価活動は, 大学教育の様々な段階で実施されるとともに, 何を評価するかという点でも多様な展開を見せている。教育政策の側からは, アセスメント・ポリシーの重要性が指摘されるとともに, 具体的なアセスメントの手法として, 学修行動調査, 学修到達度調査 (アセスメント・テスト), ループリック, 学修ポートフォリオが例示された (中央教育審議会, 2012)。また松下 (2017) は, 学修成果の評価を, 直接評価-間接評価, 量的評価-質的評価, の2次元で整理している。

他方, このような様々なアセスメント手法の試みに対して, 総合的にアセスメントの結果を把握し, 教育課程における因果関係として捉える枠組みが求められてきた (小方, 2008)。小方 (2008) は具体的なモデルとして, 学生の入学前の変数群 (学力, 家庭背景など), 大学の組織変数 (大学の教育内容, 組織), そしてアウトカム (学修成果) の間の因果関係を数量的に分析することを試みた。しかし, このような実証的な研究の蓄積は現在まで極めて不十分であり, 研究の蓄積には共通の分析モデルが構築されることが必要である (小方, 2008, pp.47-48)。

そこで本研究は, 大学教育における学修成果を可視化し評価するために, その具体的な分析の枠組みとしてのモデルを提示する。学修成果の基準となる変数は, 直接的・量的評価としての4年終了時のGPAに焦点を絞り, 一方入学時の学生の学力の指標としてプレイスメント・テストを取り上げる。さらに, 在学中の所属学科を仲介変数として取り上げることで, 3つの変数の間の因果関係をモデル化する。

モデル

本研究の目的は, 入学から卒業にいたる大学教育の過程と成果を評価するアセスメントのモデルを実証的に比較検討することである。具体的に検討されるモデルは, Astin (1993) におけるIE-Oモデルを中心に, これと比較対照される複数のモデルである。IE-OモデルにおけるI (Input) として, 大学入学時の初年次教育が開始される前の学生の準備状態を測定し

たプレイズメント・テスト（以下、PT）を取り上げる。PTは、大学入学以前に学生が身につけた学力を絶対評価するものであり、大学教育の過程への入力の一つと考えることができる。E（Environment）としては、学生が所属する学科が取り上げられる。学科による教育課程の違いは、教育環境の違いを反映するものと考えられる。さらに、O（OutputあるいはOutcome）として、卒業時（入学から4年後）のGPAの値が用いられる。4年間の学修成果の指標としては様々なものが考え得るが、学位プログラムに設定された教科・実習の総合的な成果指標としてのGPAは重要な指標の一つであると考えられるからである。

Figure 1は、本研究で検討されるIE-Oモデル（以下、モデル1）を示している。PT（I）は、3科目（英語 Eng, 国語 Jap, 数学 Math）の観測変数の背景に仮定される潜在変数である。この入力PTから出力GPA（O）へのパスが仮定されている。同時に、PTから環境変数（E）である学科（Dep）へのパスが仮定され、さらに学科から出力GPAへのパスも仮定されている。これに対して、Figure 2.は、モデル1のうち、入力PTから環境変数Depへのパスが除かれ

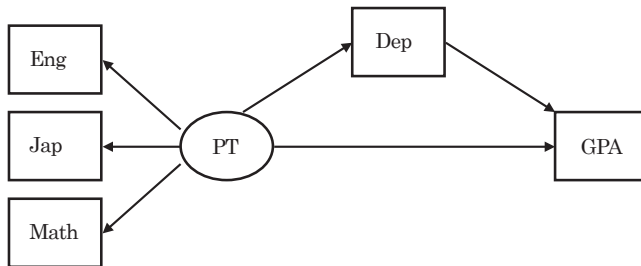


Figure 1. モデル1：入力として3つの観測変数（英語 Eng, 国語 Jap, 数学 Math）の背景に潜在変数PT（プレイズメントテスト）が仮定され、出力（成果）としてGPAが設定される。環境変数として所属学科（Dep）が設定され、PTからのパス、GPAへのパスが仮定される。Depはダミー変数として扱われる。

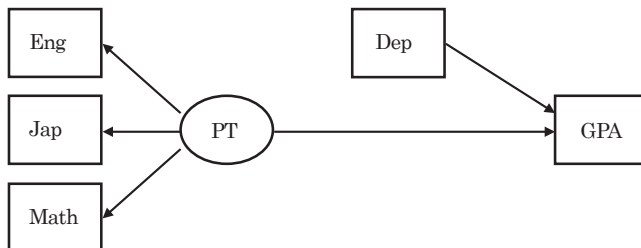


Figure 2. モデル2：モデル1のうち、入力PTから環境変数Depへのパスが除かれる。両変数の間に関連がないことが仮定される。

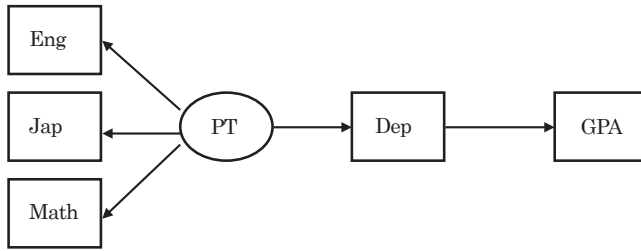


Figure 3. モデル 3：モデル 2 とは異なり，環境変数との間のみに入力からの直接のパスが仮定される。

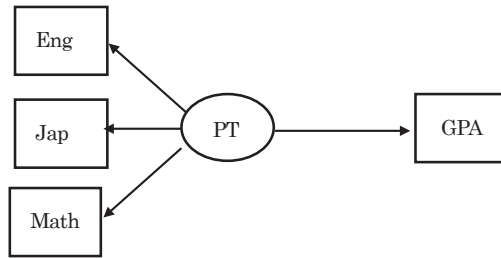


Figure 4. モデル 4：入力 PT と出力（成果）GPA のみにパスが設定され，環境変数を考慮しない。

るモデルである（モデル 2）。すなわち両変数の間に関連がないこと，環境変数としての学科が準備状態としての基礎学力に関連しないことが仮定される。Figure 3 は，モデル 2 とは異なり，入力 PT と環境変数との間のみに入力からの直接のパスが仮定されるモデルである（モデル 3）。すなわち，学科の環境が基礎学力に影響される，それによって GPA が規定されることが仮定される。最後に Figure 4 では，説明変数としての学科（Dep）はモデルから除かれ，PT と GPA の因果関係のみが仮定される（モデル 4）。以上の 4 つのモデルの比較検討から，学修成果（GPA）への学科の役割が明らかになることが予想される。

方 法

調査対象者

文系総合大学である A 女子大学の 3 学部（文学部，教育学部，社会科学系学部）8 学科の 2015 年度入学者 260 名が対象で，2018 年度の卒業年次までのデータが用いられた。データの

利用は、大学の研究倫理委員会において承認された。

調査された変数

①入学時に実施されたプレイズメントテスト（以下、PTと呼ぶ）が基礎学力の指標とされた。PTは国語、英語、数学の3科目からなっていた（以下、それぞれJap, Eng, Mathと呼ぶ）。JapとMathにはNHKエデュケーショナル制作による日本語IRT診断テスト、数学IRT診断テストが用いられた（<https://irt-test.com/>）。JapとMathは項目反応理論（Item Response Theory）に基づいてスコアが算出された（Jap: 200～800点, Math: 200～800点）。Engには、ELPA英語運用能力評価協会による大学生用の英語プレイズメント・テスト（<http://npo-elpa.org/placement/>）が用いられた。Engはリスニング、語彙・文法、リーディングの3分野各100点からなり、合計スコアが300点であった。②学生が所属する学科がカテゴリ変数として記録された（8カテゴリ：文学部3学科、教育学部2学科、生活創造学部2学科）。分析にはダミー変数が用いられた。③入学から4年後の卒業時における各学生のGPAが基準変数とされた。各科目の成績評価のAA, A, B, C, D, Nがそれぞれ4, 3, 2, 1, 0に点数化（GP）された。さらに（履修登録科目のGP×当該科目の単位数）の総和÷履修総単位数、によってGPAが学生ごとに求められた。

結 果

各変数の記述統計量をTable 1に、また観測変数間の相関行列をTable 2に示す。プレイズメント・テストの3科目の間で、英語と国語、英語と数学の間に有意な相関が認められた。また、GPAと3つのプレイズメント・テストの間にも有意な相関が認められた。

したがって3つのプレイズメント・テストの背景に潜在変数を仮定することが妥当であると判断し、この潜在変数から学科へのパスとGPAのパスを仮定し、さらに学科からGPAへのパスを加えたモデル（モデル1）を想定しパス解析を実施した（Figure 5, Table 3）。その

Table 1 プレイズメント・テスト（入学時：英語、国語、数学）、GPA（卒業時）についての基礎的な記述統計量

	英語（0～300）	国語（200～800）	数学（200～800）	GPA（0～4）
平均点	140.1	571.8	486.7	2.48
標準偏差	26.8	57.2	49.2	0.55
対象者数	258	256	256	242

注）GPA =（履修登録科目のGP×当該科目の単位数）の総和÷履修総単位数

Table 2 プレイスメント・テスト, GPA 間の相関行列

	国語	数学	GPA
英語	0.31	0.26	0.21
<i>p</i>	**	**	**
国語		0.11	0.12
<i>p</i>		<i>n.s.</i>	*
数学			0.23

(*n* = 238)

p* < .05, *p* < .01

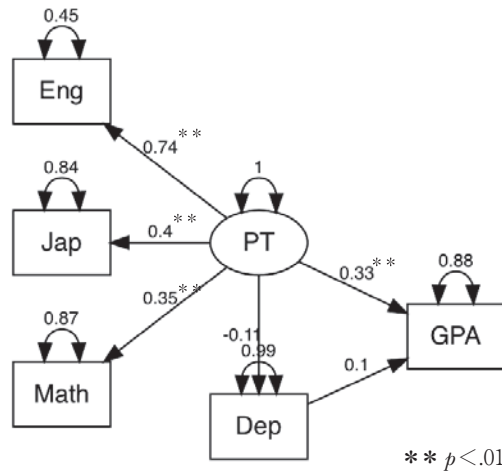


Figure 5. モデル 1 の分析結果。各観測変数上の値は残差分散である。適合度指標は $\chi^2 = 11.69$, *df* = 4, *p* < .05, *GFI* = 0.981, *AGFI* = 0.929, *RMSEA* = 0.090, 90 % *CI* [0.032, 0.152], *AIC* = 33.688, *BIC* = -10.201。GPA は学科とプレイスメント・テストによって 11.5% 説明された (*adj.R*² = 0.115)。

Table 3 検討された各モデルの適合度指標

	χ^2	<i>adj.R</i> ²	<i>GFI</i>	<i>AGFI</i>	<i>RMSEA</i>	<i>AIC</i>	<i>BIC</i>
モデル 1	*	0.1154	0.981	0.929	0.090	33.68795	-10.2011
モデル 2	*	0.1327	0.978	0.935	0.083	33.186	-14.175
モデル 3	**	0.0047	0.958	0.875	0.134	46.298	-1.063
モデル 4	<i>n.s.</i>	0.123	0.990	0.951	0.073	20.543	-6.402

注 *GFI* *AGFI* は値が大きいほど適合度が高いことを示す (< 1.0) **p* < .05, ***p* < .01
RMSEA は誤差の程度で小さいほど適合度が高いことを示す (> 0)
AIC *BIC* は値が小さいほどモデルとして優れていることを示す

結果、モデル全体としてのデータへの適合度は充分であった ($\chi^2=11.69$, $df=4$, $p<.05$, $GFI=0.981$, $AGFI=0.929$, $RMSEA=0.090$, 90% CI [0.032,0.152], $AIC=33.688$, $BIC=-10.201$)。最終的に GPA は学科とプレイスメント・テストによって 11.5%説明された ($adj.R^2=0.115$)。またプレイスメント・テストから GPA へのパス係数 (標準化係数) は 0.335, 学科から GPA へのパス係数は 0.105 であった。

さらに、モデル1からプレイスメント・テストと学科の間のパスを除いたモデル2について、同様の分析を行った (Figure 6, Table 3)。その結果、モデル2は全体としてモデル1よりも高い適合度を示した ($\chi^2=13.19$, $df=5$, $p<.05$, $GFI=0.978$, $AGFI=0.935$, $RMSEA=0.083$, 90% CI [0.029,0.139], $AIC=33.186$, $BIC=-14.175$)。最終的に GPA は学科とプレイスメント・テストによって 13.3%説明された ($adj.R^2=0.133$)。

これに対して、プレイスメント・テストから学科への直接のパスのみを仮定して、プレイスメント・テストから GPA への直接のパスを仮定しないモデル3について分析した結果 (Table 3)、データへの適合度は不十分であった ($\chi^2=26.30$, $df=5$, $p<.01$, $GFI=0.958$, $AGFI=0.875$, $RMSEA=0.134$, 90% CI [0.086,0.187], $AIC=46.298$, $BIC=-1.063$)。最終的に GPA は学科とプレイスメント・テストによって 0.5%説明された ($adj.R^2=0.005$)。

最後に、変数としての学科を含まないモデル4について検討した結果 (Table 3)、BIC以外の指標ではデータとの適合度は高い傾向であったが、BICはモデル4がモデル1、2よりも適合度が低いことを示していた。今回の分析では、学科の影響を評価することが目的であり、上

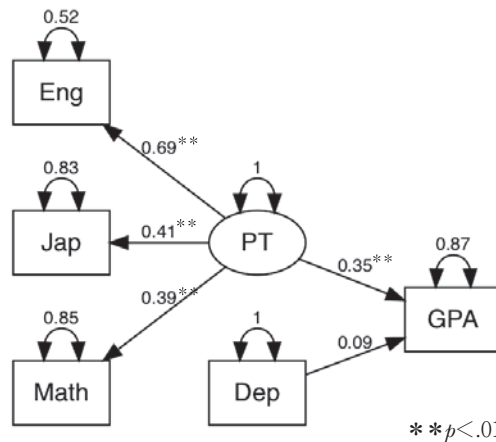


Figure 6. モデル2の分析結果。各観測変数上の値は残差分散である。適合度指標は $\chi^2=13.19$, $df=5$, $p<.05$, $GFI=0.978$, $AGFI=0.935$, $RMSEA=0.083$, 90% CI [0.029,0.139], $AIC=33.186$, $BIC=-14.175$ 。GPAは学科とプレイスメント・テストによって 13.3%説明された ($adj.R^2=0.133$)。

記モデル 1, 2, 3 の比較モデルとしてモデル 4 が設定された。GPA データの予測という点では、AIC にみられるようにモデル 4 が有効であったが、目的としたモデルの比較という観点では本来 BIC による比較が望ましい（小西・北川, 2004）。この点ではモデル 4 はモデル 1, 2 よりも適合度が低いことが示された。

考 察

本研究は、大学教育における学修成果のアセスメントのための分析モデルを提案し、具体的に検証することを目的とした。4 年間にわたる学修の成果としての GPA を基準変数として、入学時のプレイスメント・テストの結果と所属する学科、すなわち学位プログラムを説明変数とする因果モデルが比較検討された。その結果、IEO モデルにおける入力にあたる入学時の基礎学力と、出力すなわち学修成果にあたる 4 年終了時の GPA の間に直接の因果関係を仮定し、さらに学科単位の教育プログラムと GPA の間に因果関係を仮定したモデルが、学修成果を説明するモデルとして有効であることが示された（Figure 1, 2）。ただし、最も適合度の高いモデルは、入学時の基礎学力から学科の教育プログラムへの因果関係を仮定しないモデル（Figure 2, 6）であり、次いで基礎学力から学科教育プログラムへのパスを仮定したモデルであった（Figure 1, 5）。

今回の調査結果で、基礎学力と学科の教育プログラムの間に関連が認められなかったことは、学科の教育環境が学生の準備状態の程度とは関係なく設定されていたことを示している。ただし、ここで準備状態とはプレイスメント・テストで測られた学力であり、IEO モデルで一般に仮定される学生の意欲や家庭背景などの変数（山田, 2005）は考慮されていない。その結果、準備状態と学科特製の間に関連が認められなかったとも考えられる。したがって、今回の結果は IEO モデルの妥当性を否定するものではない。

つまり、ここで有効な説明を提供すると考えられたモデルは、一般的には Astin の IEO モデル（Astin, 1993）に含まれるものであり、今回の結果は学修成果の因果関係を評価するアセスメントモデルとしての IEO モデルの有効性を支持するものである。現在、IEO モデルは学生の既得情報（I）、大学の組織的特性・大学の内部機能・大学内の環境（E）、に個々の学生の努力の質を加えた Pacarella らのモデル（Pacarella & Terenzini, 2005）へと発展している。ここで学生の努力の質は、学生の個人の学修への関与（インボルブメントあるいはエンゲージメント）であるとされる。そして現在の研究の焦点は、学生の関与の重要性の認識と関与を規定する要因の検討、そして大学が関与しうる範囲の検証へと展開している（小方, 2008）。したがっ

て、本研究が仮定したモデルにさらに個々の学生の学修過程における関与の変数を加えることが今後検討される必要がある。

本研究は、大学生の学修成果を4年間にわたり追跡することにより、入学の準備段階の学力、在学する学科プログラム、4年終了時の学力の間の因果モデルを実証的に検証した。仮定されたモデルが4年間の学修の過程を説明するモデルとして学修成果の分散の13%を説明し、一定の範囲で有効であることが示された。今後、学生の学修への関与に関する変数を取り入れることにより、より有効なモデルを提示することが可能となるであろう。

引用文献

- Astin, A. W. (1993). *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*, Phenix, Arizona: ORYX Press.
- 中央教育審議会 (2012). 「新たな未来を築くための大学教育の質的転換に向けて 一生涯学び続け、主体的に考える力を育成する大学へ (答申)」.
- 中央教育審議会 (2018). 「2040年に向けた高等教育のグランドデザイン (答申)」.
- 小西 貞則・北川 源四郎 (2004). 情報量規準 (シリーズ・予測と発見の科学) 朝倉書店
- 松下 佳代 (2017). 学習成果とその可視化 高等教育研究 第20集 93-112.
- 日本私立大学連盟教育研究委員会 (2019). 私立大学における教育の質向上に関する取り組み～学修成果の可視化による大学教育の質保証～
- OECD (2008). *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*, Paris: OECD.
- Pascarella, E. T., & Terenzini, P. T. (2005). *How College Affects Students*, San Francisco, Calif: Jossey-Bass
- 小方 直幸 (2008). 学生のエンゲージメントと大学教育のアウトカム 高等教育研究 第11集 45-64.
- 山田 礼子 (2005). 一年次 (導入) 教育の日米比較 東信堂