様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成 ３０ 年　 ６ 月 ２７ 日現在

機関番号：８２６４６
研究種目：基盤研究(C)（一般）
研究期間：2015～2017
課題番号：１５Ｋ０１０３０

研究課題名（和文）指標・エビデンスの可視化による教育評価支援システム

研究課題名（英文）Educational evaluation support systems based on the visualization of indicators and evidence

研究代表者
渋井　進（Shibui, Susumu）

独立行政法人大学改革支援・学位授与機構・研究開発部・准教授

研究者番号：６０４１５９２４

交付決定額（研究期間全体）：（直接経費）　　2,800,000円

研究成果の概要（和文）：近年、大学評価に必要なデータの収集と整備が進みつつある。しかし、様々な評価の観点に対応して、どのような指標・エビデンスが有効であるかの統一がなされておらず、それに起因するコストが大学・評価機関双方にとって負担となっている。この現状を支援するために、過去に行われた評価書の分析を行うことで、指標・エビデンスの可視化をすることを目的とした。テキストの内容分析、水準判定結果の統計分析などにより、評価のプロセスの可視化がなされた。また、指標設定支援のためのチェックリストも開発した。

研究成果の概要（英文）：Recent years, the system of collecting the data to enhance the university evaluation has become become widely used in the universities. However, the absence of the standardized indicator and evidence makes it difficult to prepare for the evaluation reports of accreditation in universities and accreditation agencies. Therefore, this study examines the indicators and evidences for evaluating the universities using the text-mining approach from the past evaluation reports. The process of the evaluation was visualized by the content analysis of evaluation reports and statistical analysis of the performance levels expressed in the evaluation reports. The checklist for setting indicators to enhance the self-assessment capability of universities was also developed.

研究分野：教育工学、大学評価、認知心理学

キーワード：大学評価　教育評価　評価指標　内容分析　テキストマイニング　可視化

１．研究開始当初の背景

一般に、評価を行うにあたって、どのように評価の基準や観点を設定し、それらを判断するための根拠となるデータや指標を設定するかは、最も基本的かつ重要な問題と考えられる。しかし、そのプロセスは効率的になされていない現状がある。具体的には、測定する指標を設定する場面になると、大学の担当部署である評価室や各部局の担当部署が、経験的に考えられるものを挙げ、その中から現実的に収集可能なものを測定するという作業が行われている。

以上に関連し、近年、大学における Institutional Research(IR)の設置に見られる様に、大学評価、とりわけ教育評価に必要なデータの収集と整備が進みつつある。しかし、そのような中でも、様々な評価の観点に対応して、どのような指標・エビデンスが有効であるかの統一がなされておらず、それに起因するコストが大学・評価機関双方にとって負担となっている。

研究代表者は、ここ数年、大学改革支援・学位授与機構、鹿児島大学評価室・教育センター等の複数の機関に所属し、大学評価、特に教育評価における学習成果や、支援体制に関連する評価情報、経営情報の定義および明確化に関する研究を行ってきた。大学においては、評価基準に対しての評価指標等が明確になっていないがために、評価の都度、それらを定めるための多くの時間が割かれ、非効率になっている現状に気づいた。評価機関としては、評価が評価者の経験をもとにしたピア・レビューに依存しており、評価担当者による違いや、評価対象ごとに大きな違いが無いよう、過去の評価においてどのような根拠をもとに判定を行ったかを調べる必要性に気づき、本課題の実施に至った。

２．研究の目的

本研究では、これらの現状を支援するために、過去に行われた評価書に着目し、その内容分析を行うことで、指標・エビデンスの可視化をすることを目的とした。

評価は、大学が自己評価書を作成し、評価者がそれをもとに判断して、評価結果報告書として判定する流れで行われる。この中の、自己評価書と、評価結果報告書を、教育評価に関する評価基準・観点ごとに、手作業による内容分析及び、テキストマイニング手法を用いて分析し、頻出して用いた指標等を明らかにした。

さらに、自己評価書と評価結果報告書の間のギャップも、評価の水準判定データをもとに分析することで、大学と評価機関のギャップが明らかになると予想された。

以上より得られたデータを、多変量解析等の統計的手法を用いて可視化を行うことで、現在一般的に用いられている、専門家によるアートの部分が多かった、自己評価に基づいたピア・レビューのプロセスが「見える化」され、評価の省力化、効率化に向けた基礎的なデータとなることが予想された。

これらのデータを大学、評価機関に公開することで、今後の大学内での評価の遂行や評価制度の設計の参考として利用を促進することを目的とした。

また、評価指標を可視化するだけではなく、新たに大学が指標を設定する時に自らの指標設定の妥当性について判断可能な指針の作成も、支援の一環として作成した。

３．研究の方法

以上の目的のため、大別して３つの方法により、研究を行った。以下、それぞれに記述を行う。

（１）　評価書の内容分析による指標・エビデンスの可視化

過去に大学機関別認証評価を受けた大学の自己評価書及び、評価結果報告書のテキストデータを、手作業およびテキストマイニング手法によって自動化することで、頻出する指標・エビデンスや、統計分析により、用語の共起分析等を行い、グラフィカルに表示することで、可視化を行った。

分析対象とした評価の基準・観点は、「単位制度の実質化」、「学習成果」、「教育の内部質保証システム」、「成績評価の厳格性」などが代表的なものであった。

（２）　自己評価書と評価結果報告書に記載された大学・評価者間での水準判定のギャップの分析

国立大学法人等の第２期中期目標期間における教育研究の状況の評価結果のデータを用いた。各大学の提出した実績報告書および、評価結果の中から抽出した学部・研究科等の現況分析結果において、評価対象の大学等による自己評価及び機構による評価結果の水準判定データを分析対象とした。評価は学部・研究科別に教育及び研究に関する２つの評価が実施されており、教育の評価に関する項目は、「教育活動の状況」、「教育成果の状況」の２つの分析項目があり、その下に観点として、「教育実施体制」、「教育内容・方法」、「学業の成果」、「進路・就職の状況」の 4 つが観点として設定されており、以下、教育に関してはこの４つを評価の観点のデータとして用いた。

研究の評価に関する分析項目は、「研究活動の状況」、「研究成果の状況」の２つであった。その下に、「研究活動の状況」、「共同利用・共同研究の実施状況」、「研究成果の状況」の観点が設定されていた。その中で、「共同利用・共同研究の実施状況」に関しては、共同利用機関等の少数が対象であったことから除外し、「研究活動の状況」および「研究成果の状況」の２つを評価の観点のデータとして用いた。

評価の水準はいずれにおいても、4 段階（4：期待される水準を大きく上回る、3：期待される水準を上回る、2：期待される水準にある（標準）、1：期待される水準を下回る）による段階判定であった。これらを点数化して、自己評価、他者評価の別、評価の観点、学系を要因とした 3 要因の分散分析のデータをもとにグラフ化し、可視化を行った。

（３） 指標妥当性のチェックリストの作成

まず、評価指標にかかる大学現場の状況を把握するために、具体的な設定にまつわるプロセス設定の難しさ等について先進的と考えらえる２大学を選び、予備的ヒアリング調査を実施した。

それを踏まえ、指標の妥当性に関する先行研究の文献レビューを行った。その結果、教育心理学における心理測定および、政策評価の特に目標管理・達成度評価、の２つの分野で中心的に扱われてきたことがわかり、両者における基礎的な概念を整理した。

心理測定においては質問紙における尺度構成の妥当性、政策評価においては評価指標の妥当性という違いはある。しかし、これらを比較すると、異なった文脈の下に妥当性の検討がなされており表現に違いはあるのだが、妥当性を構成する基本的要素自体は共通している点が多い。そこで、それらの共通点を整理し、妥当性概念を整理した。それをもとに、大学評価の文脈でのチェックリストとその解説を作成し、さらに事例を付け加えた教材を作成した。

４．研究成果

上記３つの方法に対応し、得られた結果の中から代表的なものを以下に示す。

（１） 評価書の内容分析による指標・エビデンスの可視化

ここでは、いくつか行った分析の中から、「単位制度の実質化」について内容分析を行った結果を示す。

図 1a, 1b に、大学評価・学位授与機構が実施した大学機関別認証評価の 2005 年度から 2011 年度までの 7 年間（第１サイクル）および 2012 年度から 2014 年度までの 3 年間（第２サイクル）を合わせた 10 年間における評価報告書の、学士課程および大学院課程それぞれの「単位制度の実質化」に関する評価項目の記載データを分析対象とし、サイクル間での指標・エビデンスの出現頻度に関して比較を行った結果を示す。

また、図 2 にサイクル別および課程別の 4 カテゴリと指標・エビデンスをカテゴリとし、出現頻度を集計したクロス表を元に、コレスポンデンス分析を用いて 2 次元空間上にマッピングを行って、相対的な布置関係を表示した結果を示す。

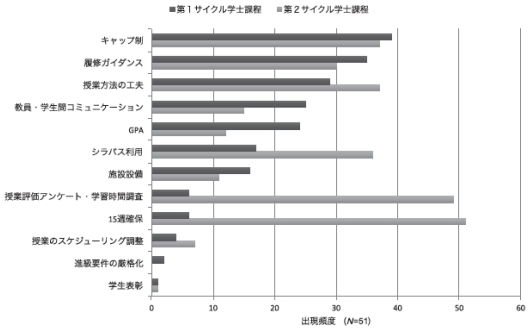

図１a 学士課程におけるサイクル別の指標・エビデンスの出現頻度

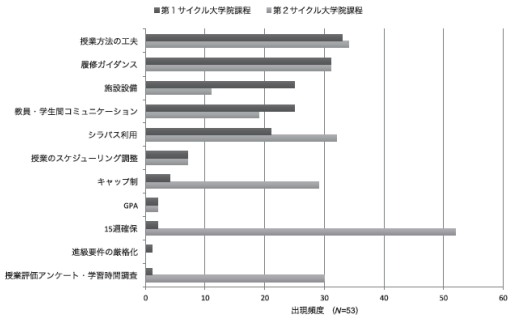

図１b 大学院課程におけるサイクル別の指標・エビデンスの出現頻度

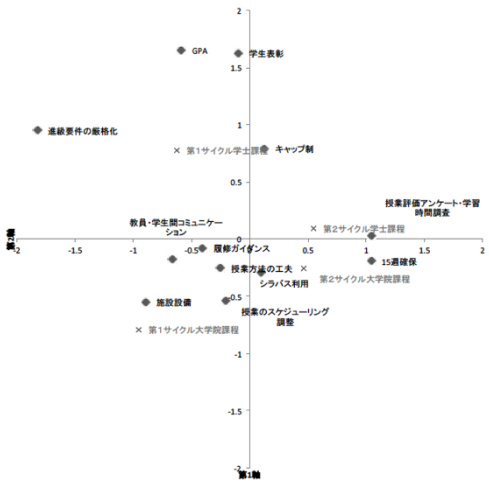

図 2 コレスポンデンス分析によるサイクル間、課程別での指標・エビデンスの出現傾向

以上より、サイクル別に記載の差異が見られたキャップ制、GPA、シラバス利用などの記載から、中教審を通した政策の影響が大学に強く浸透していることが解釈できた。その一方で、教員・学生間コミュニケーション、履修ガイダンス、授業方法の工夫、などの大学が普遍的に掲げてきた指標がコレスポンデンス分析の空間の中間付近に位置することから解釈でき、指標が明確となるとともに、それらの特性についても可視化された。

（２） 自己評価書と評価結果報告書に記載された大学・評価者間での水準判定のギャップの分析

教育、研究のいずれにおいても、被験者内要因として自己評価、他者評価の別（以下、評価の主体と呼ぶ）、評価の観点、被験者間要因として学系を要因とした３要因の分散分析を適用した。結果は以下の通り。

①　教育に関して（図 3,4）

３要因すべてにおいて主効果が有意であった。

・評価の主体（F（1, 805）= 1521, $partial\ \eta^2$ = .654, p < .01）

・評価の観点（F（3, 2415）= 79.61, $partial\ \eta^2$ = .090, p < .01）

・学系（F（9, 805）= 4.130, $partial\ \eta^2$ = .044, p < .01）

交互作用に関しては、評価の観点と学系のみに有意差が見られ、

・評価の観点×学系（F（27, 2415）= 3.038, $partial\ \eta^2$ = .033, p < .01）

その他の交互作用には有意差は見られなかった。

・評価の主体×評価の観点（F（3, 2415）= .987, $partial\ \eta^2$ = .001, p = .398）

・評価の主体×学系（F（9, 805）= 1.664, $partial\ \eta^2$ = .018, p = .094）

・評価の主体×評価の観点×学系、F（27, 2415）= 1.459, $partial\ \eta^2$ = .016, p = .060）

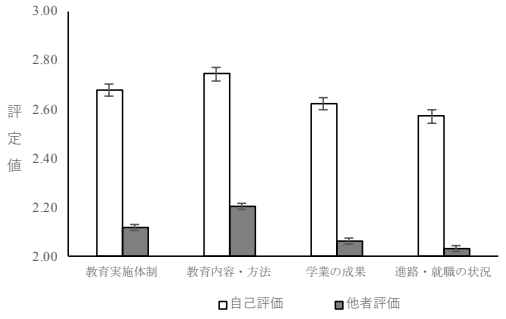

図３　評価の主体別の評価の観点における教育水準判定結果の評定値

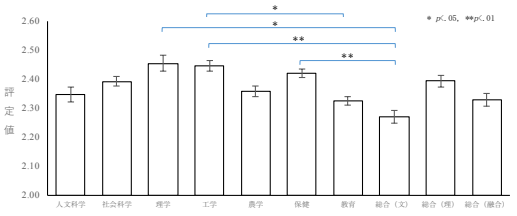

図４　学系別の教育水準判定結果の評定値

②研究に関して（図 5,6）

３要因すべてにおいて主効果が有意であった。

・評価の主体（F（1, 581）= 251.3, $partial\ \eta^2$ = .302, p < .01）

・評価の観点（F（1, 581）= 10.39, $partial\ \eta^2$ = .018, p < .01）

・学系（F（10, 581）= 6.107, $partial\ \eta^2$ = .095, p < .01）

交互作用に関しては、評価の主体と評価の観点、評価の主体と評価の観点と学系、に有意差が見られ、

・評価の主体×評価の観点（F（1, 581）= 12.80, $partial\ \eta^2$ = .022, p <.01）

・評価の主体×評価の観点×学系、F（10, 581）= 2.286, $partial\ \eta^2$ = .038, p < .05）

その他の交互作用には有意差は見られなかった。

・評価の主体×学系（F（10, 581）= 1.785, $partial\ \eta^2$ = .030, p = .060）

・評価の観点×学系（F（10, 581）= 1.782, $partial\ \eta^2$ = .030, p = .061）

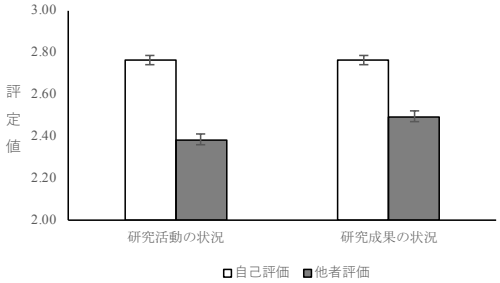

図５　評価の主体別の評価の観点における研究水準判定結果の評定値

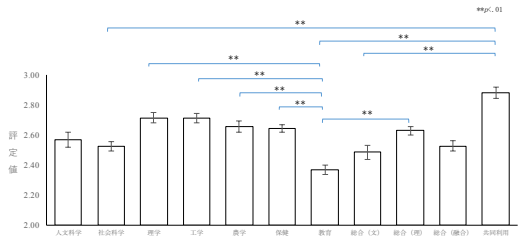

図６　学系別の研究水準判定結果の評定値（自己評価、他者評価の平均値）

全体として教育も研究も、自己評価の方が他者評価よりも高く評定される傾向がみられた。原因を、以下に考察する。

まず、理由の一つとして、評価結果が運営費交付金に反映されるという国立大学法人評価の性質を考えると、必然的に高い評価を得るために大学がプレッシャーを受けることとなる。それゆえ、戦略的な視点から、自己評価を高く見積もり、他者評価が高くなるようにアピールする傾向が影響した可能性が考えられる。

また、心理的なバイアスとして、社会心理学的な立場からは、自己評価において一般に望ましい特性は平均的他者よりも高く見積もられる傾向として、平均以上効果が知られている。認知バイアスが自己評価に影響を及ぼしている可能性もある。これは、文化差や、評価の文脈による違いはあるだろうが、大学

評価に置き換えて考えると、他大学よりも自大学の評価を高く見積もってしまう傾向があるかもしれない。

このように、自己評価と他者評価（評価結果）のギャップが明らかになり、その背後にある要因について考察がなされた。

（３） 指標妥当性のチェックリストの作成

完成したチェックリスト（表 1）と教材の例は以下の通りであった（図７）。

表１ 指標妥当性のチェックリストと解説

区分	基準	説明
妥当性	目的との適合性	指標が，計画の進捗や目指す成果を適切に反映しているか。
	調査対象・結果への影響	指標設定の結果，意図しない悪影響を及ぼすものではないか。
	信頼性	誰がいつ測定しても，同じ事象や状態からは同じ測定結果が得られるか。
	理解可能性	指標の意味が，明確でわかりやすく，誤解が生じないか。
	包括性・非重複性	計画の重要な側面が，もらさず指標によってカバーされているか。
		指標間に重複がなく，各指標は異なる側面を計測しているか。
実用面	意思決定者への有用性	指標が，執行部等の意思決定者に対して，有益な知見を提供してくれているか。
	計測可能性	指標となるデータは収集可能か。
	収集の適時性	有用なタイミングで，指標の計測値は入手可能か。
	データ収集のコスト	データを収集するための費用は大きすぎないか。
	操作可能性	指標の計測値は，都合良く操作して変更可能なものではないか。

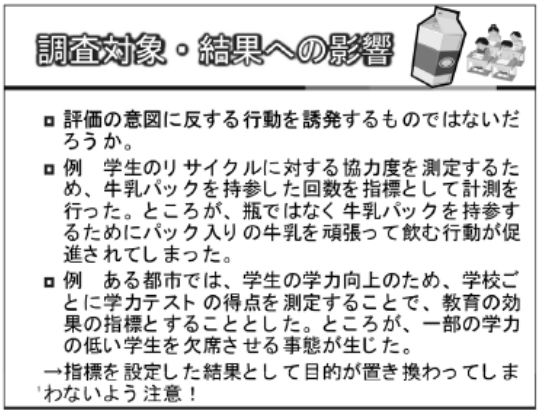

図７ 教材の例（調査対象・結果への影響に関する妥当性の解説）

これらの結果は、論文等として公表することで、大学・評価機関の実務者を含めた関係者に幅広く周知することで、普及を図った。また、評価機関や大学内での研究会等でも報告を行った。

５．主な発表論文等

〔雑誌論文〕（計 4 件）

① 渋井進，野田文香（2018）．評価書分析による「単位制度の実質化」に係る指標・エビデンスの可視化，*大学評価・学位研究*，19，39-55.（査読あり）
http://id.nii.ac.jp/1107/00000434/

② 渋井進，浅野茂，橋本貴充，小野宏， 小野達也，山﨑その，田中弥生（2017）．自己評価力向上支援のための評価指標設定に関するチェックリストの開発，*大学評価・学位研究*，18 21-36.（査読あり）
http://id.nii.ac.jp/1107/00000422/

③ SHIBUI Susumu, TAKAHASHI Nozomi, NODA Ayaka (2017). A Longitudinal Study of Effectiveness, Impact, and Challenges in the Japanese Quality Assurance System, *International Journal of Institutional Research and Management*, 1, 83-102.（査読あり）
http://www.iaiai.org/journals/index.php/IJIRM/article/view/118

④ 野田文香，渋井進（2016）．「単位制度の実質化」と大学機関別認証評価，*大学評価・学位研究*，17，20-33.（査読あり）
http://id.nii.ac.jp/1107/00000397/

〔学会発表〕（計 18 件）

① SHIBUI Susumu, NAKATO Emi (2018). “An exploratory analysis of the factor affecting education and research performance levels in Japanese national universities based on the evaluation reports”, 12th annual International Technology, Education and Development Conference.

② 渋井進，坂口菊恵（2018）．「教育・研究水準の判定結果に影響する要因の検討 ～ 自己評価と評価結果の関係を中心に ～」，電子情報通信学会ヒューマンコミュニケーション基礎研究会.

③ 渋井進（2017）．「大学評価・IR と心理学」，日本基礎心理学会第 36 回大会.

④ NODA Ayaka, SHIBUI Susumu (2017). “Quality Assurance in University Student Learning”, The 9th Asian Conference on Education.

⑤ 渋井進，高橋望（2017）．「自己評価と評価結果のギャップをもたらす要因の検討-国立大学法人評価のデータを用いて-」，日本心理学会第 81 回大会.

⑥ 渋井進（2017）．「認証評価における内部質保証の考え方 ～内部質保証を評価する方法を中心に～」，高等教育質保証学会第 7 回大会.

⑦ 渋井進，高橋望（2017）．「大学評価における自己評価と評価結果の関係 ～国立大学法人評価の現況分析結果をもとに ～」，電子情報通信学会ヒューマン情報処理研究会.

⑧ SHIBUI Susumu, TAKAHASHI Nozomi (2017). “An Exploratory Data Analysis of the Indicators and Evidence for Measuring the Internal Quality Assurance System”, INQAAHE Conference 2017.

⑨ 渋井進 (2017). 「認証評価における内部質保証の考え方 ～内部質保証を評価する方法を中心に～」, 高等教育質保証学会第 7 回大会.

⑩ SHIBUI Susumu, TAKAHASHI Nozomi (2017). “Visualization of Indicators for Measuring the Internal Quality Assurance System in Japanese Universities”, The 15th Hawaii International Conference on Education.

⑪ 渋井進, 高橋望 (2016). 「教育の内部質保証システムに対する大学の認識の多様性」, 電子情報通信学会ＨＣＧシンポジウム 2016.

⑫ 高橋望, 渋井進, 野田文香 (2016). 「機関別認証評価の大学アンケート分析 から見えるもの ― 第１サイクルと第２サイクルの比較を中心に ―」, 第 5 回大学 情報・機関調査研究集会.

⑬ 渋井進, 田中弥生 (2016). 「大学評価支援へ向けた指標設定のチェックリストの開発」, 日本評価学会春季第 13 回全国大会.

⑭ SHIBUI Susumu, TAKAHASHI Nozomi, NODA Ayaka (2016). “Visualization of the cognitive dimensions for evaluating universities by means of the content analysis of evaluation reports” International Meeting of the Psychonomic Society 2016.

⑮ 高橋望, 渋井進, 野田文香 (2016). 「認証評価を通じた教育改善の取組状況の分析-第２サイクルにおける改善機能に着目して-」, 大学教育改革フォーラム in 東海 2016.

⑯ 渋井進, 高橋望, 野田文香 (2015). 「大学評価支援に向けた評価指標・エビデンスの可視化」, 電子情報通信学会ＨＣＧシンポジウム 2015.

⑰ 渋井進, 野田文香 (2015). 「『教育の内部質保証システム』に関する評価書の内容分析」, 日本心理学会第 79 回大会.

⑱ SHIBUI Susumu, NODA Ayaka (2015). “Content analysis of the evaluation reports to find the evidence for measuring the effectiveness of the credit hour system” The 14th European Congress of Psychology.

〔図書〕
なし

〔産業財産権〕
なし

〔その他〕

① アウトリーチ活動として、2017 年 2 月 17 日に九州地区大学 IR 機構勉強会にて「大学評価と指標・ エビデンス」というタイトルで、依頼講演を行った。
http://www.u-ryukyu.ac.jp/info/IR2017030802/

６．研究組織
(1)研究代表者
渋井　進 (SHIBUI SUSUMU)
大学改革支援・学位授与機構・研究開発部・准教授
研究者番号：60415924

(2)研究分担者
野田　文香 (NODA AYAKA)
大学改革支援・学位授与機構・研究開発部・准教授
研究者番号：20513104

(3)連携研究者
なし

(4)研究協力者
なし