

ループリックの法学部教育への展開

—法学部学生研究発表大会における 実践例と統計的検証—

石 上 敬 子
喜 入 暁
大 島 一 悟

1	はじめに	1
2	法学部学生研究発表大会の概要	2
	(1) 経法大・「法学部学生研究発表大会」	
	(2) COVID-19 と 2020 年度の大会運営	
3	ループリックの開発	3
	(1) 審査基準へのループリック導入の経緯	
	(2) 2019 年度：「評価観点」の再設定	
	(3) 2020 年度：「評価基準」の具体化	
4	ループリック評価の実践と検証	13
	(1) 目的と方法	
	(2) 結果	
	(3) 考察	
5	おわりに	22

1 はじめに

本稿は、本学法学部の「学生研究発表大会」（全学年参加の学生プレゼンテーション大会）で用いた「ループリック」について、その内容と開発過程、および検証結果をまとめたものである。

「ループリック」は、学修目標到達度の評価基準として、大学教育の場でも徐々に広がりつつある⁽¹⁾。しかし、法学分野では公表された例は少なく、とりわけ《法学部》における教育、つまり法学分野・政治学分野における教育についてみると、政治学分野での例はほとんど知られていない⁽²⁾。

研究ノート

本稿の実践例は、学生のプレゼン大会における《順位付け》のための評価基準という、ルーブリックの本来の趣旨からすればやや特殊な場面のものである^③。しかし、法学・政治学分野における評価基準の一例として、特に法学・政治学分野では例の少ない、評価データの統計分析を通じた信頼性の検証結果もあわせて示したい。

2 法学部学生研究発表大会の概要

(1) 経法大・「法学部学生研究発表大会」

本学の「法学部学生研究発表大会」（以下、大会とする）とは、法学部生全学年による、「法学・政治学等の観点から現代社会の諸課題を取り扱った研究」の成果を発表するイベントである。大会は法学部だけでなく、本学の全ての学部ごとに企画され、11月末の同日に一斉に開催される。法学部の大会の中核は、パワーポイントなどを用いた口頭発表であり、法学部教員らによる審査員が質疑応答を経てこれを評価し、優れた発表には表彰状等が授与される。

ただし、この大会はいわゆる「卒論発表会」とは異なる。出場資格は法学部生であることに限られ、学年の制限はなく、個人・グループのどちらでも良い（グループの人数制限もなく、他学年と組むことも可能）。4年生の出場はむしろ少なく、1年生も出場し、入賞する例がある。一方、大会は各学年の「演習」と密接なかかわりがある。大会のグループ結成、発表準備などは、演習中で行われることが多く、また多くの演習が、出場者以外には大会観戦レポートを課している。このため、幅広く全学年から、観戦者も含めて多数の学生が参加する場となっている^④。

近年の実施状況は次の通りである。出場組数、出場者数は2018年度以降減少しているが、《量より質》を目指して、発表時間を拡大したこと（10分→15分）の影響と考えられている。

② 阪経法論 84(21.3)

表 1 大会実施状況

	出場組数	出場者数	参加者総数
2020 年度	20 組	41 名	確認不能（後述）
2019 年度	24 組	60 名	646 名
2018 年度	35 組	84 名	630 名
2017 年度	50 組	123 名	710 名
2016 年度	53 組	162 名	696 名

(2) COVID-19 と 2020 年度の大会運営

もっとも、2020 年度は COVID-19 の影響によって、大会企画の大きな変更を余儀なくされた。

COVID-19 による年度当初からの全面的な遠隔授業の展開の中、法学部では既に 4 月には、大会の 11 月開催自体を決定事項とし、学生向けにも公示した。これは「遠隔授業の今こそ演習の学びに目標を」との趣旨に基づくものであったが、その後の感染状況も踏まえつつ、《大会感》を可能な限り維持しうる方法を検討した結果、「無観客試合（発表者と審査員のみが対面、観客は zoom で遠隔リアルタイム観戦）」という初の形態が採用されることとなった。

しかし、それ以外の部分についてはできるだけ前年を踏襲して、大会運営の安定的な実現を目指した。その中で唯一、意識的に取り組んだのが、次に述べる「審査基準」の改善である。

3 ルーブリックの開発

(1) 審査基準へのルーブリック導入の経緯

審査基準は毎年、見直し・改善を重ねてきたが、2019 年度からルーブリック導入を試みることになった⁽⁵⁾。

ルーブリックとは一般に、「学習到達目標の達成度を視覚化するための図表」であり、より簡単には、「ある課題について、できるようになってもらいたい特定の事柄を配置するための道具」とされる。基本的な要素は、課題、評価尺度、評価観点、評価基準であり、次のような形式をとる。

表2 基本的なルーブリックの表

課題			
	評価尺度 1	評価尺度 2	評価尺度 3
評価観点 1	評価基準 1-1	評価基準 1-2	評価基準 1-3
評価観点 2	評価基準 2-1	評価基準 2-2	評価基準 1-2
評価観点 3	評価基準 3-1	評価基準 3-2	評価基準 1-3

「評価観点」は「課題」を細分化したものであり、評価の視点を示す。「評価尺度」はいわゆる点数である。そして、ルーブリックで最も重要な点は、「評価基準」において、学生が到達すべき目標や、その過程においてつまづきやすい点を、文章によって具体的に記述することである⁽⁶⁾。

これを大会の審査基準として採用することで筆者らが狙ったのは、次の3点である。

- a) 審査基準の客観化・明確化
- b) 発表の質的向上
- c) 観戦者に対する教育的効果

ルーブリックの意義は一般に、①「評価者・被評価者の認識の共有」、②「複数の評価者による評価の標準化」にあるとされる⁽⁷⁾。筆者らがまず意識したのは、大会における審査基準、つまり順位付けの評価基準という目的のゆえに、②評価者（審査員）による評価のばらつきを小さくすることであった（= a)）。一方、ルーブリックは通常、複数の評価者が利用することを常に想定するものではなく、その意味で第一の意義は、①評価の基準を学生に対し明示し、学生と共有することとされる。このことは大会においては、発表者との関係では、高い評価を受ける発表とは何かを具体的に示すことを通じて発表の質的向上が期待され（= b)）、また観戦者との関係では、発表をただ傍観するのではなく、ルーブリックを手がかりに自らも批評者となることを通じて、自分の普段の学びのあり方も見直すきっかけを得ることが期待された（= c)）。

(2) 2019年度：「評価観点」の再設定

2019年度は、評価観点の再設定から着手した。手がかりとしては、それ

④ 阪経法論 84(21.3)

までの審査基準と、本学法学部のディプロマ・ポリシー (DP)⁽⁸⁾ をベースとして、日本学術会議「大学教育の分野別質保証のための教育課程編成上の参照基準」の法学分野、政治学分野に関する部分 (=「参照基準法学 2012」「参照基準政治学 2014」) を参照した。これら参照基準はもちろん、ルーブリックを念頭においたものではないが、各専門分野の学修を通じて獲得すべき素養や、そのための評価方法を示す点で、ルーブリックで言う「評価観点」の参考になると考えたことによる⁽⁹⁾。以下、やや長くなるが、関連する部分を引用する。

参考 1：参照基準法学 2012

○法学を学ぶ学生が身に付けるべき基本的素養 (iv、9-12 頁)

- a. 日本国憲法を中心とする国家規範の構造の理解、特に国民主権、三権分立、基本的人権の尊重
- b. 制定法の理解と、その基礎にある様々な価値観や法治主義、適正手続、罪刑法定主義、所有権の保護、契約自由等の法原理の理解
- c. 法を運用する様々な機関の理解
- d. 法的判断の持つ特殊性の理解
- e. 法的判断過程における「説得」の重要性の理解
- f. 様々な考え方の分析能力とコンセンサスを得るための調整能力の涵養
- g. グローバル化に伴う様々な法的問題とその基礎にある文化の多様性の理解

○評価の視点 (19-20 頁) ※傍線筆者

「法学は、様々な異なった価値観、利害に応じた意見や主張が存在することを認識の前提にしつつ、それらの考察にはその基礎にまで立ち返って考察・検討することを重視している。そのため、学修の評価にあたっては、様々な異なった見解やその背後にある様々な価値観などを十分理解した上で、それらを適切に調整しつつ適切な論拠を採出し、具体的な結論に至る論理を構築することが重視される。様々な学説や条文、裁判例、具体的事件等の情報を裏づけにしつつ、論理的に議論を組み立て、ある一定の結論を導く能力が、評価されるのである。」

「個々の科目における評価のあり方は、教育目標、知識のレベル、教育方法などにより異なっている。……教育方法別に評価のあり方の大まかな特徴を述べれば、次のとおりである。」

【以下、一部抜粋】

「(c) 論文作成 卒業論文やゼミ論文等の場合には、その内容は課題やアプローチによって極めて多様であり、一律の評価尺度や達成すべき水準の指標は存在しない。もちろん、先行研究や内外の文献や判例を適切に調べているかどうか、それらを活用して論理的で一貫性を持った議論が展開されているかどうか、実証や論証の手続きが厳密であるかどうか、学術論文としてのルールや作法が遵守されているかどうかなど、卒業論文やゼミ論文等として評価すべきポイントは多岐にわたる。しかしながら、何をどのぐらい重視して評価するかについ

研究ノート

ては、それぞれが、教育課程全体の中で果たす役割に照らして適切に判断されるべきであり、またそのことについて教員間の理解の共有が図られるべきことは、他の教育方法と同様である。」

「(e) 実習・現場教育、(f) プレゼンテーション能力の向上等では、学生の行動や発言を教員が注意深く観察することが、評価にとって重要な手がかりとなる。……授業場面での観察であれ、事後的なふり返りであれ、それを評価するにあたっては一律の評価尺度や達成すべき水準の指標は存在せず、どの要素をどう評価していくかは、深い知識を有する評価者の高度な評価・判断能力に依存することになるが、これらを明確に意識して行うことが重要である。」

「法学教育においては、このような多様な評価が組み合わせられて用いられる。学修の最後に行われる総括的評価では、学理的な知識・技能の修得やその活用に重点を置いて総括的な評価がなされる場合もあろうし、学んだ知識・技能を実践的・応用的に使いこなす力量に重点を置いて総括的な評価がなされる場合もあるであろう。それぞれの大学における目標の設定の仕方や、それぞれの授業のねらいによって、多様な方法と多様な評価の仕方が採用されることは、多様な特長を有する人材を社会に送り出すという意味で、尊重されなければならない。」

参考2：参照基準政治学 2014

○政治学を学ぶ学生が身に付けるべき基本的素養 (ii, 9-12 頁)

「政治現象を認識する学問である政治学を学ぶ意義の中心は、政治に関してより合理的に考察し判断できるようになることである。

その際、中心となるのは、政治家や官僚に必要とされる能力よりも、市民として政治を観察し、それに積極的に関わってゆくための思考力と判断力を身に付けることである。よりよい市民のあり方（シティズンシップ・市民性）の涵養が、大学での政治学教育の中心課題である。」

○評価の視点 (13-14 頁) ※傍線筆者

「……政治学教育の使命はあくまでも、政治に直接関わる職業に従事するわけではない、市民のための思考力と判断力の涵養にある。」

「たとえば、広汎な現実の中から政治に関わる問題をどうやって発見し、いかにして解決方法を論理的に導き出すか。政治思想のテキストをいかに着実に読み解くか。実証データをどうやって集め、そこからいかにして合理的な結論を引き出すか。先行研究を吟味した上で、新しい研究主題をどこに見定めるか。そういった学問的な訓練を積みながら政治現象について学ぶことで、自らの思考をどれほど柔軟に展開できるようになったかを検証するような評価方法が必要である。」

これらを見ると、評価の観点が一定程度までは示されているが、共通して強調されているのは評価方法の多様性であり、評価方法の一般化は避けられている。

一方、他に手がかりとなりうる評価基準として、早稲田大学公共経営大

学院における次のものがある。

参考3：早稲田大学公共経営大学院リサーチペーパー（政策提言論文）の評価基準（近江2016：120-128頁）

- 1 公共経営領域に関する課題を取り上げ、テーマ設定が適切であること
- 2 修得した分析手法に則って研究がされていること
- 3 矛盾なく一貫した論旨が展開されていること
- 4 研究内容の記述や展開が説得的であること
- 5 実行可能性その他の評価などを検討した経過が認められること
- 6 必要最低語数を満たし、規則に則った書式であること
- 7 博士後期課程への進学をめざすものは、独創性があること
- 8 自らが設定した問題について、自らの調査・分析に基づき回答を試みていること

以上を踏まえ、評価観点の大枠については、次のような案が示された。

表3 評価観点（案）

評価観点	審査の観点	関連する
①問題提起	具体的かつ説得的な理由に基づいて問題提起がなされているか ・テーマを選定した理由（問題意識）が明確・具体的に示されているか ・テーマの意義について、説得的な説明がなされているか	DP2
②調査手法	的確な知識を基礎とし、多面的な視点を有する研究がなされているか ・法と政治に関する基本的専門知識を正しく理解しているか ・複数の学説・判例など、多様な視点への配慮がなされているか	DP1 DP3
③分析手法	合理的な裏付けのある論理的な分析がなされているか ・テーマの説明・分析が、獲得した知見に適切に基礎づけられているか ・結論を導く論理展開に飛躍がないか	DP4
④発表技術	受け手のことを考えた発表がなされているか ・口頭発表：声の大きさ・スピード、資料（ppt等）、時間配分（減点事由） ・レジュメ：段落等の全体構成、文章表現の洗練度、参考文献	DP4
⑤総合評価	質疑応答も踏まえ、総合的に優れた研究と評価できるか ・基本評価基準＝DP2 ✓ 国民主権、権力分立、基本的人権の尊重など基本的事項を踏まえているか ✓ 様々な価値観や法治主義、適正手続、罪刑法定主義、所有権の保護、契約自由等の法原理を理解しているか ✓ 法的判断の持つ特殊性を踏まえ、「説得」の重要性を理解しているか ✓ 様々な考え方の分析、コンセンサスを得るための調整がなされているか ✓ グローバル化に伴う法的問題と文化の多様性を理解しているか	DP2

①～④は、2018年までの枠組のものを踏襲しつつ、「早稲田院・評価基準」を参照して審査の整理し、またDPとの関係を確認したものである。一方、⑤は、従来の質疑応答評価に加えて、総合的な評価項目を加える目

研究ノート

的で新設した。前述した「評価基準」で示された評価要素のうち特に法学分野から、大会の趣旨に沿うものを引用した（政治学分野は法学分野ほど細かく示されていなかったため、今回は引用していない）。

ところで、同年度はこれ以降の検討時間が十分確保できず、評価基準などの詳細は大会当日にようやく完成・公開できたにとどまった。このため、同年度のループリックおよび実施状況等については詳論しない。

(3) 2020年度：「評価基準」の具体化

前年度の結果を受け、2020年度は改めてループリックの完成に向けて、特に法学部の教育に関して既に公表されているループリックを参照して検討を行った。ただし、政治学分野ではループリックはほとんど知られておらず、法学分野で次の山下2018（表4）、吉永2019（表5）が公表されているのみであった。

山下2018は、法的事例問題に対する答案（「法律文書」）の評価基準であり、このため法学三段論法を評価観点として採用した上で、「学生が陥りやすいミス」を詳細に列挙したものである。さらに、「全体としては良く書けているが細かい部分が気になる答案や、個々の論点は書けているが全体としてつながっていない答案」があることを考慮して、「答案全体としての評価」をS～Dで冒頭に示している点や、「全体の講評や評価基準になる記述の補足」のためのコメント欄を設けている点に特徴がある（ただし後者は、評価基準の充実化に伴って使用頻度が下がっているという）。

吉永2019は、法律系専門科目の講義レジュメとして、指定の文献をパワーポイントでまとめたものに対する、自己評価のための基準である⁽¹⁰⁾。a)はそうした場面の特殊性ゆえの項目であり、またc)は、指定された文献の丸写しを防ぐことが強調されている。一方、b)（情報の内容）は大会の評価基準として参考になり、またd)・e)は、パワーポイント発表に対する評価の基準として参考になる。

以上を踏まえ、表6のようなループリックを完成させた（カッコ内は配点、満点36点）。

⑧ 阪経法論 84(21.3)

表 4 山下 2018 (107 頁を一部整形・省略)

法律文書評価シート

Ver.5.0

S	A	B	C	D
<p>さん</p> <p>大要よくできている</p> <p>ルールの提示：複数の条文解釈の可能性を自らの主張を考慮するなど、ルールの選択を支える一般論自己の目的にあわせて意識的にその根拠を示せるか？</p> <p>・条文の文言からその場で、法理のない解説論を示すなど、法律が向士でも通用するようなルールを提示を行う能力が認められる。</p> <p>・事実関係の細かい点にも目を配り、事実の再解題なども行い、具体的な事実・関係、事実に照らして、要件を指摘して、要件に該当することを論証できている。</p> <p>結論を導く論理：最終結論まで三段論法で組み合わせて提示する高度な能力が認められる。</p> <p>三段論法で組み合わせて提示するこ とができるか？</p> <p>・法的三段論法を組み合わせるまでの論理過程を明確に提示することができ る。</p> <p>・相手方の反論を予測して再反論を提示する。 ・例外ルールを提示する。</p>	<p>よくできている</p> <p>ルールについて、要件と効果を正確に理解し、明確に提示できる。</p> <p>・ルールの背後にある根拠を示しつつ、法源から認得力のある形でルールを導ける。</p> <p>・提示したルールとの関係で、重要な事実を指摘して、要件に該当することを論証できている。</p>	<p>改善の余地がある (1)</p> <p>ルールを選けているが、要件や効果の理解が一部不正確である。 ・ルールが必要に複雑なっている。 ・ルールの根拠を十分示していない。 ・ルールの根拠が曖昧である。 ・学説の提示に熱心でルールがわかりにく い。 ・あてはめと混ざってルールがわかりにく い。</p> <p>重要な事実を指摘できているが、 ・ルールと事実の対応の論証が不十分であ る。 ・ルールと事実の対応関係が不自然である。 ・ルールと事実の対応関係がわかりにく い。</p> <p>法的三段論法を組み合わせて提示するこ とができるか？</p> <p>・法的三段論法を組み合わせるまでの論理過程を明確に提示することができ る。</p> <p>・段落わけなど読みやすさへの配慮が足 りない。 ・論証に不要な文章が議論を冗長にして いる。 ・論理がみられる。 ・サブルールに分解できる部分が無視し ている。</p>	<p>改善の余地がある (2)</p> <p>十分なルールの提示とはいえない。 ・要件や効果の理解がかなり不正確である。 ・条文とルールの対応関係が明確でない。 ・ルールの根拠が不適切、不正確。 ・ルールの提示に不要な議論が記されている。 ・条文不提示。</p> <p>十分な論証とはいえない。 ・ルールの事実の対応がまったく取れていない。 ・重要な事実の指摘が欠けている。 ・事実の指摘が短すぎて評価できない。 ・事実の指摘が整理されておらず評価できな い。 ・提示されたルールに問題があり、あてはめの能 力評価不能。</p> <p>十分な論証とはいえない。 ・結論に至る論理過程が不十分である。 ・示かれている内容に正面から答えていない。 ・結論までの論理に破綻がある。 ・論理過程の複雑すぎて認得力が低い。 ・論証の能力を疑わせる部分がある。 ・守備の欠落など、誤用など文のつながりが不明 確である。 ・ルールとあてはめは問題があり、論証の能力評 価不能。</p>	<p>大幅に改善の余地がある</p> <p>要件を満たすと効果が生じるとい意味でのルールといえるものが示されていない。 ・ルールが法律に根拠を持つものである必要を理解していない可能性がある。</p> <p>・ルールの中の要件を、事実と照らして判断するということの意味を理解していない可能性がある。</p> <p>・自分の導きたい結論までの論理過程を示すことができている。 ・自分の導きたい結論までの論理過程を示すことができている。</p>
<p>コメント</p>				

表5 吉永 2019 (4頁を一部整形)

民法Ⅱ自己チェック用ルーブリック (評価項目表)			
	レベルS	レベル1	レベル0
a) 文献ルール	教科書(講義計画に掲載)に加えてガイダンスレジュメの文献リスト上野文献2冊(六法や法律用語辞典などは数に加えない)を参考文献とし、参照した頁の範囲を特定して、最終スライドに掲げている。	指示した文献の数を満たしていない、または、参照した頁の範囲を記載していない。 (レジュメを2点減点する) *厳格に採点する(版の違いにも注意!)。教科書+文献リスト上の文献2冊で合計3冊とすること。	指示した文献を1つもあげていない、あるいはそもそも参考文献を最終スライドに掲げていない。 (レジュメを0点とする)
b) 情報の内容	豆知識の羅列に終わらず、法律を学ぶ上で中心となる情報(要件・効果、制度趣旨、具体例)を核としてレジュメが組み立てられ、その量が十分である。	法律を学ぶ上で中心となる情報を核としてレジュメが組み立てられているが、所々に余計な情報、欠けている情報、不正確な情報がある。	豆知識の羅列に終わっている、あるいは情報量が大きく不足している。
c) 情報の提示の仕方	1文の情報量が少なくなるように情報を適切に切り分けており、さらに並べた情報相互のつながりを示すなど、レジュメ全体の関係性を示す整理がされている。	情報を適切に切り分けて提示してはいるものの、情報の羅列・箇条書きにとどまっており、整理がされていない。あるいは論理のつながり、情報の順序におかしいところがある。	教科書の説明文を(多少手を加えて)写し取るにとどまり、1文に複数の情報を押し込むなど、情報の切り分けができていない(とりわけ記述内容を理解できていないとの印象を与えるものを含む)。
d) 誤字等	誤字や日本語としておかしい表現がない。		誤字や日本語としておかしい表現がある (個数分だけレジュメ点を減点する)。
e) プレゼンテーション資料としての見やすさへの配慮	字の大きさ・色・書体、行間、見出しの階層化・テキストのインデントなど、資料の見やすさに配慮されている。		資料の中に見にくさを感じさせる箇所がある。

初めに前提として、評価対象となる発表には法学、政治学双方に関するものが含まれることから、山下 2018 のように法学に特化した基準としない方針を確認した。一方で、《法学・政治学には共通するが、他の分野では見られない》評価の観点を示すことに苦慮した結果、分野独自の評価の観点を「総合点」という項目に集約する一方で(3(2)参照)、それ以外の項目は一般的な評価観点にとどまるものとなった。ただしその際、山下 2018 でも見られたように《ルーブリックで表現しきれない評価》を尊重する趣旨から、「総合点」は配点を敢えて10点と多く設定し、評価基準も具体的に示さずに柔軟な判断の余地を残した。このため「総合点」はルーブリックとしての特徴を欠くものとなったため、以下ではこの部分を「総合点評価」と呼び、それ以外の部分を「ルーブリック評価」と呼んで区別する。

表 6・大会ルーブリック (2020 年度)

2020 年法学部学生研究発表大会・審査基準ルーブリック

●会場 A B C D ●発表順 1 2 3 4 5 ●審査員名

	評価	満点 (2~1点)	0 点
問題提起 (4)	問題提起が曖昧ではなく明確に示されている。		問題提起がなされていない、または曖昧である。
背景 (3)	問題提起に至った背景が、主観的または曖昧でなく、問題設定と乖離していない (問題提起が必然的に導かれる)。		問題提起に至った背景が、主観的または曖昧でなく、問題設定と乖離していない場合も含む。
情報源 (3)	発表者の主張以外の情報 (判例・学説・事実/事例) のうち一部に、情報源 (参考文献など) が示されている。 ※ Wikipedia や個人ブログなど、信ぴょう性に乏しい情報は評価対象としない (情報源が示されていないものとして扱う)。		発表者の主張以外の情報 (判例・学説・事実/事例) のすべてに、情報源 (参考文献など) が示されていない。
視点 (3)	複数の視点から引用した情報のすべてが、論の構成に寄与している (情報回士とのつながりが明確である) / それ以外の情報が不十分な議論になり得る。		情報の複数の視点から引用されているが、不要な情報の視点から引用されているが、 <u>意図</u> によらずとまっている。
主張・主張分析・主張アプトアップ (10)	問題提起と対応した主張/結論が、曖昧ではなく明確に示されている。 主張/結論が、論理の飛躍や強引な結論導出がなく、明らかな根拠 (事実/事例を含む) を基に論理的に導かれている。		主張/結論が示されていない、または曖昧である。 主張/結論が問題提起と対応していない。 主張/結論の根拠 (事実/事例を含む) が示されていない。
批判的検証 (3)	異論/反論を的確に取り込み、それらに対して説得力のある応答が示されている。		異論/反論に触れられていない。
残された課題 (3)	残された課題が明確に挙げられている (今回の研究で、何ができ、何ができなかったのかを、明示できている)。		残された課題に触れられていない。
発表技術 (6)	聴衆を意識した、聞き取りやすいプレゼンテーションである。		声の大きさとスピードの両方に難がある。
プレゼン資料 (3)	パワーポイントの特性を活かした資料になっている。 (文字量は必要最低限、図・写真など視覚的要素を促す工夫)。		パワーポイントの特性を活かす工夫が見られない。 (図がない、文章をそのまま貼り付け、字が小さいなど)
その他 (減点事由)	発表時間が 14 分 30 秒 - 15 分 30 秒でない。 (± 30 秒につき 1 点減点、ただし 16 分以上は強制終了し 2 点減点) 誤字脱字 (4 種類以上の場合、マイナス 1 点)		
総合点 (10)	質疑応答を踏まえ、総合的に優れた研究と評価できるか。 ※ 以下は判断基準の例です。各観点を踏まえた上で、総合的に評価してください。 ・ 国民権、権力分立、基本的人権の尊重など基本的事項を踏まえているか ・ 様々な価値観や法治主義、憲正手続、罪刑法定主義、所有権の保護、契約自由等の法原理を理解しているか ・ 法的判断の持つ特殊性を踏まえ、「説得」の重要性を理解しているか ・ 様々な考え方の分析、コンセンサスを得るための調整がなされているか ・ クローバル化に伴う法的問題と文化の多様性を理解しているか		

研究ノート

ルーブリック評価の各項目の趣旨は次の通りである。

評価観点は前年度に検討したものを前提に（表現は若干修正した）、項目をさらに細分化した。評価尺度を満点か0点かの2値評価とした項目（「問題提起」「主張」）は、当然に満点となることが期待される要素であり、その他の評価尺度は0点から3点の4段階評価として、教員（審査員）がルーブリックを使用した経験が乏しいことを踏まえて、若干の裁量の幅（減点2～1点）を残した。

評価観点の「問題提起」の項目「背景」では、単に興味を持った、目についたので調べたという発表（＝背景が主観的・あいまい）を低く評価することを意図した。

評価観点「調査」の項目「情報源」では、情報源選別の観点を明示しつつ（wikipedia等に関する記述）、パワーポイント発表ではあるが、できるだけ参考文献一覧ではなく、参照情報が都度付記されるよう、「すべてに…明示されている」との表現を採用した。項目「視点」では、吉永2018のc)を参考に、情報の羅列や、一面的な主張をなぞることのないよう示唆した。「……情報のすべてが、論の構成に寄与している」という表現では、引用された情報の関係が論理的に整理されることを示唆しつつ、議論の単なる紹介を排除すること（主張・結論を導くために有益な情報だけが挙げられること）を狙っている。

評価観点の「分析・主張」の項目「論理性」では、《論理的であること》について、飛躍していないこと、強引でないことに加え、根拠の説得力（「明らかな根拠」という観点も含めて表現した。項目「批判的検証」は、反対説に対する応答の適切さを問うものであり、「説得力のある応答がなされている」という表現は具体性を若干欠くものの、減点事由における《的確にとらえられているか》（「正しくとらえていない」という表現において、反対説を正しく理解し、噛み合った応答をすることの重要性を強調している（藁人形論法の排除）。さらに、項目「残された課題」は、教育的意図をもって新設した。これまでの大会ではほとんどの発表が、この点に触れていなかったという実態を踏まえて、研究で扱った部分と扱わなかった部分の区別、今後の展望などを、発表者が意識することを促そうとした。

評価観点の「発表技術」では、配点合計を全体の6分の1に抑えることで、パフォーマンス大会ではないとのメッセージを打ち出した。項目「プレゼンスキル」では、原稿の棒読みとならないよう示唆し（減点事由参照）、項目「プレゼン資料」では、吉永 2018 の e) を参照しつつ、パワーポイントの細かな機能に凝りすぎないように、適切な活用を促している。項目「減点」は、吉永 2018 の d) を参照しつつ、発表を聞きながらの審査となるため誤字脱字を正確に数え上げるのが難しいことを考慮して、4 個以上で一律減点とした。

4 ルーブリック評価の実践と統計的検証

(1) 目的と方法

こうして作成したルーブリックを、2020 年度大会では発表の審査基準として本格的に導入した。その目的は次の3点である（3(1) 参照）。

- a) 審査基準の客観化・明確化
- b) 発表の質的向上
- c) 観戦者に対する教育的効果

ただし a) について補足すると、その具体的目的は「評価者（教員）間における評価のばらつきが小さくなること」だと述べたが、今回のルーブリックの用途は「大会における順位付けの評価基準」であるため、ルーブリックの本来の趣旨（＝すべての評価対象が万点評価に収束することが理想⁽¹¹⁾）とは異なって、「発表間における評価のばらつきが大きくなること」（＝発表ごとのレベルの違いを捉えられている）も要請されることになる。前者を「ルーブリックの評価基準としての信頼性」（以下「a1」とする）、後者を「ルーブリックの審査基準としての有効性」（以下「a2」とする）とし、以下の統計分析では特に、この2点の検証を目的とする。

ルーブリックは完成後、10月に学部教授会で報告して承認されたのちに、学生に対しては大会出場受付開始（大会1ヶ月前）と同時に公表した。ただし、審査にルーブリックを用いることは、6月の大会要項発表時点で公表しており、参考として昨年のルーブリックを示して、準備の手がかりとするよう周知していた。

研究ノート

教員（審査員）に対しては、11月半ば（大会1週間前）に改めて説明会を行い、ループリックの趣旨、各項目の意味を一通り説明した上で、昨年の大会動画を用いて模擬審査を実施し、質疑応答を行った。

大会当日、出場組数は20組であった（2(1)参照）。これを4会場に5組ずつ振り分け、また審査員は各会場に3名ずつ割り当てて、各審査員は会場5組の発表の審査を一貫して担当した。各会場への発表者と審査員の割り振りは、発表テーマと教員の専門分野等を考慮して大会実行委員が決定した。

審査の対象は、発表および直後に行われる質疑と、発表で用いたパワーポイントデータ（参考資料として配布）とした。審査の時間は、発表15分とその後の質疑が終わるまでとし、審査員にはループリックを印刷した用紙を配布して、結果を記入した用紙を発表終了ごとに回収した（＝事後の審査修正は不可）。表彰は、全会場の最高得点を「最優秀賞」、第2位・3位を同率の「優秀賞」とし、各会場では第1位から第3位までを「会場1席」～「3席」とした。

一方、聴衆として観戦する学生に対しても、それぞれが評価者となって大会に参加するよう、演習等を通じて積極的な周知を行った（演習によっては評価の提出を授業課題の一部としている）。ただし、学生による評価は表彰とは切り離し、あくまで観戦者の大会へのかかわりを具体化するためのものと位置付けた。また、今年は観戦者はオンライン参加（zoom）としたため、学生による評価の集計にはWebサービス（Microsoft Forms）を利用した（なお、webサービス上ではトラップ項目を挿入し、努力の最小限化（satisfice⁽¹²⁾）が検出されたデータは排除した）。学生による評価の対象は、発表と質疑だけとした（パワーポイントデータは配布していない）。また、webサービスの入力には期限を設けなかったため、学生は時間の制約なく評価を行っている。

(2) 結果⁽¹³⁾

発表に対する大会の公式評価である、教員（審査員）3名の評価を平均した値（以下「発表ごと教員評価」とする）は、全体では表7および図1

⑭ 阪経法論 84(21.3)

(1)の通りとなった。標準偏差は評価のばらつきを表すが、変動係数(相対標準偏差)はばらつきの度合いの比較に用いられる。一方、教員による全ての評価、および学生による全ての評価(以下「教員全評価」「学生全評価」とする)は、表8、図1(2)(3)の通りであった。

表7 発表ごと教員評価

※()は満点に対する得点率

	平均値	中央値	最大値 ^a	最小値	標準偏差	変動係数
発表ごと教員評価 (発表20組)	22.4 (62.1%)	23.0 (63.9%)	27.7 (76.9%)	17.3 (48.1%)	3.38	0.15

^a最大値=最優秀賞。

表8 教員全評価、学生全評価

※()は満点に対する得点率

	平均値	中央値	最大値	最小値	標準偏差	変動係数
教員全評価(発表20組 ×教員審査員3名)	22.4 (62.1%)	23 (63.9%)	32 (88.9%)	7 (19.4%)	5.12	0.23
学生全評価 (評価学生60名) ⁽¹⁴⁾	28.2 (78.3%)	29 (80.1%)	36 (100.0%)	11 (30.6%)	5.87	0.21

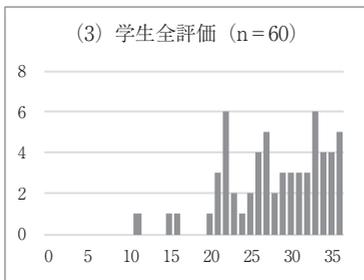
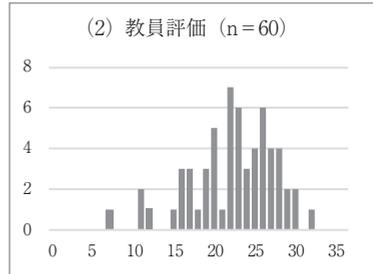
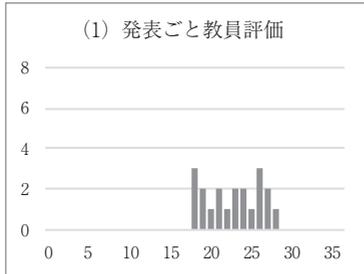


図1 評価の分布

研究ノート

「教員全評価」と比べると、ばらつきは「学生全評価」のほうがやや大きい（標準偏差 5.87）。また評価も「学生全評価」のほうが総じて高く、満点評価が5件存在している⁽¹⁵⁾。

さらに、教員による評価を評価項目ごとにみると、表9、図2・3の通りであった。ここでは教員（審査員）間における評価の一致度検証のため、級内相関係数（intra-class correlation：ICC）も算出した（詳細は（3））。

表9 教員による評価（評価項目ごと） ※（ ）は満点に対する得点率

	平均値	標準偏差		変動係数		ICC	審査員間 一致数
		（全評価）	（発表ごと）	（全評価）	（発表ごと）		
問題提起（0-1） ^a	55	91.67				0.14	16
背景（0-3）	2.3	0.86	0.62	0.38	0.28	0.32	5
情報源（0-3）	2.3	0.73	0.39	0.32	0.17	-0.05	0
視点（0-3）	2.0	0.84	0.45	0.42	0.22	-0.07	2
主張（0-1） ^a	49	81.67				0.24	13
論理性（0-3）	1.9	0.79	0.49	0.41	0.26	0.11	1
批判的検証（0-3）	1.5	0.93	0.47	0.62	0.31	-0.06	2
残された課題（0-3）	1.6	1.05	0.69	0.66	0.44	0.18	0
プレゼンスキル（0-3）	2.1	0.81	0.58	0.38	0.28	0.31	4
プレゼン資料（0-3）	2.3	0.75	0.43	0.32	0.18	0.02	3
ループリック評価（0-26）	16.9 (64.9%)	4.38	2.99	0.26	0.18	0.22	0
総合点評価（0-10）	5.5 (54.7%)	1.48	1.00	0.27	0.18	0.24	0

^a 2値（0か1）での評価であるため、平均値の列には1の度数、標準偏差の列には1の割合を記載している。

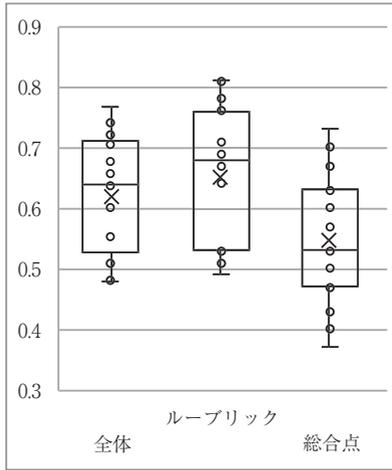


図2 評価の分布（評価項目ごと得点率）

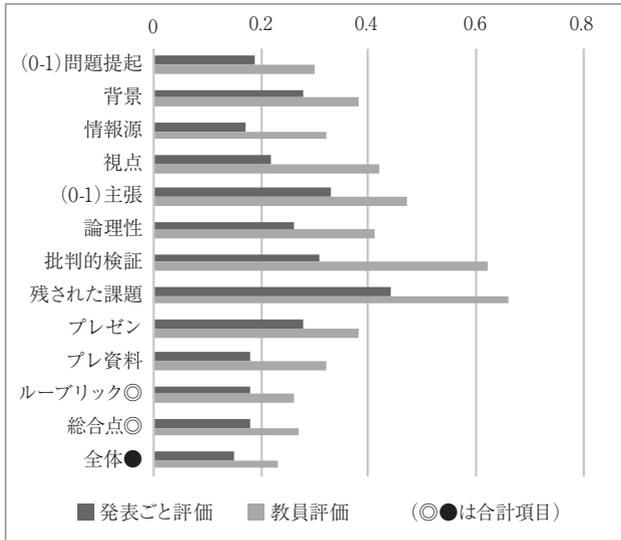


図3 評価のばらつき（評価項目ごと変動係数）

研究ノート

発表に対する評価を項目ごとにみると、ルーブリック評価の平均値は16.9（得点率 64.9%）であり、最大値は 24（得点率 92.3%）、最小値は 6（得点率 23.1%）であった。対して総合点評価は、平均値は 5.5（得点率 54.7%）であり、最大値は 8（得点率 80.0%）、最小値は 1（得点率 10.0%）であった。さらに、ルーブリック評価の項目ごとにみると、2 値評価（0-1）とした「問題提起」「主張」はいずれも満点の割合が多く（「問題提起」= 91.7%、「主張」= 81.7%）、他の項目では「背景」「情報源」「プレゼン資料」が平均値が最も高く（2.3/3）、平均値が低い項目は「批判的検証」（1.5/3）、次いで「残された課題」（1.6/3）のであった。評価観点の大分類でみると、最も評価が高いのは「発表技術」（=「プレゼンスキル」「プレゼン資料」）であり（評価観点内平均値 2.2/3）、また「調査」（=「情報源」「視点」）もわずかな差にとどまる（評価観点内平均値 2.15/3）。「問題提起」（=「問題提起」「背景」）も、2 値評価を含むため直接の比較はできないが評価は高かったと言えよう。対して「分析・主張」は、2 値評価でない「論理性」「批判的検証」「残された課題」が総じて低かった（平均値 1.67/3）。

さらに、以上のデータを元に、ルーブリックの評価基準としての信頼性、ルーブリック評価と総合点評価との関連性について分析を加えた。

ルーブリックの信頼性は、一般に、評価者間における評価の一致によって示されると考えられており（=一致度が高いほど信頼性が認められる）、級内相関係数（ICC）がそのための第一の指標として用いられる（= ICC が 1 に近いほど信頼性が認められる）⁽¹⁶⁾。

これを教員全評価でみたところ、ルーブリック評価とその各項目、および総合点評価のほぼ全てにおいて、一致度が非常に低いことが示された。級内相関係数は値の大きいものでも、「背景」（ICC=0.32）、「プレゼンスキル」（ICC=0.31）にとどまり、ルーブリック評価と総合点評価を比較しても、大差なく低い結果となった（ICC ルーブリック評価=0.22、ICC 総合点評価=0.24）。評価の一致数を具体的に見ると、教員審査員 3 名の評価が一致した数は、「背景」でも発表 20 組のうち 5 組（25%）、ジェネラルスキルであった「プレゼンスキル」でも 4 組（20%）であり、それ以下との差もわずかである。とりわけ「情報源」「残された課題」は、全く一致しなかったことが明らか

⑱ 阪経法論 84(21.3)

になった。ただし、「問題提起」「主張」のICCの低さは、これらが2値評価であり、かつ評価が全体に「1」に偏っていたことに由来するものであり、⁽¹⁷⁾「1」での一致を具体的に見ると、「問題提起」では発表20組のうち16組(80%)、「主張」では13組(65%)であった。さらに総合点評価では、完全に一致した組はなかったものの、教員3名の評価の差が1以内にとどまったものが5組存在している(ただし評価の差の最大は4であり2組存在した)。

続いて、ルーブリックによる評価と総合点評価との関連性を、相関分析によって検証した。

教員全評価を対象とした分析の結果、統計的に有意な正の相関関係が示された($r=0.38$, $p=0.003$)⁽¹⁸⁾。ただし、相関の強さは中程度にとどまり、ルーブリック評価と総合点評価にはある程度の乖離が生じている可能性が示唆された。一方、学生全評価を対象とした分析の結果、強い正の相関関係が示された($r = 0.82$, $p < 0.001$)。ただし、表8・図2に関連して述べた通り、学生による評価はどのような発表に対しても高い傾向がみられたため、この意味を一般化することには慎重であるべきと考えられる。

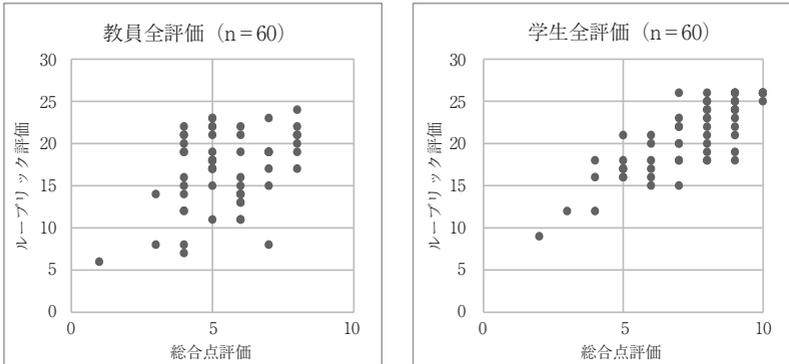


図4 ルーブリックによる評価と総合点評価との相関

(3) 考察

ルーブリックの評価基準としての信頼性 (a1) については、本研究では評価者間における評価のばらつき (一致度) を、教員全評価における級内相関係数などによって分析した (表9)。その結果、ルーブリック評価とその各項目、および総合点評価のほぼ全てにおいて、一致度が非常に低いことが示された。評価の一致度につき、ルーブリック評価と総合点評価との差もほとんどみられなかったことは、今回のルーブリック評価は、評価基準を具体的に示さない柔軟な判断基準であった総合点評価よりも、信頼性を有するものとなり得なかった可能性を示唆している。この点、本研究では実施過程において、大会前に審査練習会を行い、審査員教員12名を含む15名以上が同一対象を評価する場があったが、データの収集・分析はせず、質疑と意見交換にとどめていた。今後はむしろ積極的に、そうした機会を早期から何度か設けてデータの検証を行い、改善・推敲を重ねるべきと考えられる。

これに対し、ルーブリックの審査基準としての有効性 (a2) について、発表間における評価のばらつきを、変動係数によって分析した (表7～9)。変動係数はばらつきの比較に用いられうる数値であり、単独でばらつきの大小を評することはできないが、過年度についてみると2019年度審査基準の変動係数は0.18、2018年度は0.19であった。2020年度の変動係数は0.15 (発表ごと教員評価) であり、これらを下回ってはいるが、その差はわずかとも評しうる (ルーブリック評価、総合点評価の数値も近い)。

ところで、ルーブリック評価と総合点評価との関連性については、教員による評価を対象とした相関分析の結果、統計的に有意な関連は示されたものの、想定されていた強い関連は示されなかった。この原因としては先に述べた通り、ルーブリック評価の信頼性が不十分だったからとも考えられるが、ルーブリックについては評価基準の明確性の問題だけでなく、評価観点が網羅的になっていなかった可能性も考えられる。すなわち、開発過程でも述べたように、今回のルーブリックでは、「法学・政治学の研究発表に特化した専門的評価観点」を示すことができず、その受け皿となる可能性を念頭に総合点を残した経緯がある。この点、事後に行ったルーブリ

ック評価に関する意見では、「総合点評価はルーブリックの評価項目となっていない部分を補填するための項目と捉えていた」といったものがあり、相関が低いことによってこそ、ルーブリック評価と相互補完して評価観点を補っていた可能性が考えられよう。

次に、ルーブリックの項目ごとに検討する。

2 値評価の「問題提起 (0-1)」「主張 (0-1)」は、評価者間では高い一致度が確認される一方 (a1)、満点 (= 1) での一致の割合が高く、発表間のばらつきは小さい結果となった (a2)。このことは審査基準としての有効性に疑義を生じうるが、これが現実として意味するものは、多くの発表がこの基準をクリアしている、ということであり、教育的観点 (b) からはむしろ望ましい状態である。配点が抑えられている限りはバランスも保たれることから、今後は他の項目でも同様に、発表間のばらつきが小さくなった場合には配点を下げるという対応が考えられよう。

一方、「情報源 (0-3)」「残された課題 (0-3)」では、評価者間の一致は全くみられなかったが、「情報源」のばらつきが相対的に小さいのに対し、「残された課題」のばらつきは相対的に大きくなった。この点「情報源」は、15分という発表時間内で精査するのが困難だった (審査に混乱を生じ一致が少なくなった) ことが考えられ、審査の方法自体を再検討すべき可能性がある。これに対し「残された課題」は、「批判的視点 (0-3)」にも同様の傾向がみられた。また0点と3点という最大の乖離が生じたケースも、「残された課題」で2/20組、「視点 (0-3)」で3/20組がみられた。「残された課題」は教育的意図から新設された項目だったが、まずは教員間でその意義を議論し、共通理解を形成することが求められるだろう。

5 おわりに

今回のルーブリックは、評価者間の一致度が十分に示されなかったことから、信頼性については不十分であったと言わざるを得ない。今回の分析で問題点が明らかになった項目を含めて今後も検証を重ね、また教員間での議論の場を可能な限り設けて、改善を図ってゆくべきであろう。総合点

研究ノート

評価はルーブリック評価との関係では、乖離していたことに一定の意義を見出しうる可能性があるが、評価者間の一致が低い限り、信頼性のある評価基準となり得ない。今後の課題は、総合点評価考慮要素をゆくことであろう。

一方、ルーブリックによる教育効果という点で注目すべきは、今年の発表の入賞組は、ルーブリックの評価観点を意識したとみられるものが多かったことである。評価基準に照らして水準が高かったかはともかく、事前にルーブリックを公表して参照するよう周知したことは、学生に研究調査の進め方に手がかりを与える効果があったことがうかがわれる。

このほか、観戦者に対する教育的効果などについては今後、年単位で比較を行い、調査を追加してゆく必要がある。ルーブリックの開発はまだ緒についたばかりであるが、今後もさらなる向上を目指したい。

引用文献

参照基準解説 2017

日本学術会議・大学教育の分野別質保証推進委員会、「大学教育の分野別質保証のための教育課程編成上の参照基準について(解説)」、<http://www.scj.go.jp/ja/member/iinkai/daigakuhosyo/pdf/kaisetsu.pdf>、(2021. 01. 06)

参照基準政治学 2014

日本学術会議・大学教育の分野別質保証推進委員会 政治学分野の参照基準検討分科会、「大学教育の分野別質保証のための教育課程編成上の参照基準政治学分野」(平成26年(2014年)9月10日)、<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-h140910.pdf>、(2021. 01. 06)

参照基準法学 2012

日本学術会議・大学教育の分野別質保証推進委員会 法学分野の参照基準検討分科会、「大学教育の分野別質保証のための教育課程編成上の参照基準：法学分野」(平成24年(2012年)11月30日)、<http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-22-h166-2.pdf>、(2021. 01. 06)

中教審答申 2012

中央教育審議会、「新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～(答申)」(平成24年(2012)8月28日)、https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1325047.htm、(2021. 01. 11)

スティーブンス 2014

ダネル・スティーブンスほか（佐藤浩章ほか訳）『大学教員のためのルーブリック評価入門』（2014年、玉川大学出版部）

近江 2016 近江幸治『学術論文の作法』（2016年、成文堂）

會田ほか 2018

會田信子ほか「大学生の協同学習を評価するルーブリック試作版の考案：看護学生を対象とした信頼性・妥当性の検討」看護科学研究16巻2号33-52頁（2018年）

沖 2014 沖裕貴「大学におけるルーブリック評価導入の実際：公平で客観的かつ厳格な成績評価を目指して」立命館高等教育研究14号71頁以下（2014年）

貝谷ほか 2017

貝谷敏子ほか「看護演習科目へのルーブリック導入の効果・ルーブリック評価の信頼性と妥当性の検討」札幌市立大学研究論文集11巻1号3-11頁（2017年）

対馬 2010 対馬栄輝「信頼性指標としての級内相関係数」弘前大学 医学部 保健学科 統計学資料2、<https://personal.hs.hirosaki-u.ac.jp/pteiki/research/stat/icc.pdf>

三浦・小林 2015

三浦麻子・小林哲郎「オンライン調査モニタの Satisfice はいかに実証的知見を毀損するか」社会心理学研究31号120-127頁（2015年）

森田 2014 森田果『実証分析入門 データから「因果関係」を読み解く作法』（2014年、日本評論社）

山下 2018 山下純司「ルーブリックの法学教育への応用」学習院大学法学会雑誌53巻2号103-115頁（2018年）

吉永 2019 吉永一行「学生の発表を中心にした大講義科目の運営手法：『橋本メソッド』の一応用例として」高等教育フォーラム（京都産業大学）9号113-118頁（2019年）

オーサーシップ貢献（CRediT authorship contribution）

石上敬子：創案（Conceptualization）、研究の実施（Investigation）、データのまとめ（Data Curation）、体系的な解析の実行（Formal Analysis）、執筆-草案の作成（Writing - original draft preparation）、執筆-草案の見直しおよび再構成（Writing : review & editing）。

喜入暁：創案（Conceptualization）、研究の実施（Investigation）、データのまとめ（Data Curation）、体系的な解析の実行（Formal Analysis）、執筆-草案の見直しおよび再構成（Writing : review & editing）。

大島一悟：研究指導 (Supervision)、プロジェクトの統括 (Project Administration)。

-
- (1) たとえば中教審答申 2012では、大学教育で「速やかに取り組むことが求められる事項」として、成績評価基準の明確化においてルーブリック等を用いることが示唆されている (20頁)。また 沖 2014 では、導入の具体的な手順等が、日本の初等中等教育やアメリカ大学教育での活用事例を元に示されている。
 - (2) 法学分野での先例は、山下 2018、吉永 2019 のみとみられ、政治学分野での先例はみつけられなかった。本事例ではこの山下教授、吉永教授の両名にも、ルーブリックの内容について相談し、ご意見を仰いでいる。
 - (3) 後にも述べるが、ルーブリックは本来、「学生が達すべき目標を満点評価において示すこと」を意義とし、「学生の多数が優劣なく満点評価を得ること」を理想とする。このため、大会などでの《順位付け (差別化)》の基準としては、本質的には適さないと考えられている。
 - (4) 学内で他に研究成果を発表する場としては、2月をメ切とする「法学部学生顕彰論文」があり、こちらは卒業論文など4年次生を中心とする、個人の研究発表の場となっている。
 - (5) ルーブリックの概要をまとめた文献としてはスティーブンス 2014が著名であり、法学分野の文献では 山下 2018 が詳しい。本稿ではこれらを参照しつつ、簡潔な説明にとどめている。
 - (6) 中教審答申 2012の用語集 (39頁) では、ルーブリックは「評価水準である『尺度』と、尺度を満たした場合の『特徴の記述』で構成される」と説明されている。
 - (7) 中教審答申 2012の用語集 (39頁)。
 - (8) DPの内容は次の通りである。
 1. 自由、平等、民主主義などの価値原理を基礎とする、法と政治に関する基本的専門知識を体系的に理解している。
 2. 個人から国家に至るまで当事者間で発生する諸課題を理解し、法的知識・技能をもとにして、多様な価値観や利害関係に適した解決策を考える力を身につけている。
 3. グローバル化する現代社会の一員として、他者と協調・協働できる能力を身につけている。
 4. 論理的な思考力と豊かな表現力とともに幅広い教養および実践感覚を身につけている。

- (9) 「参照基準」とは、日本学術会議が2008年に文部科学省から委託を受けたことを契機に作成されたものであり、大学教育の質保証のために、学士課程の専門教育の教育課程編成において参照されるべき基本的な考え方を、分野ごとに示したものである。各分野共通の枠組として、各学問分野の定義・特性を確認した上で、各分野の教育を通じて身につけることを目指すべき「基本的な素養」と、それを修得するための「学習方法」と「学習成果の評価方法」についての基本的な考え方が示されている（参照基準解説 2017）。
- (10) 吉永教授は、200名を超える規模の大講義を、受講生がレジュメを作成して講師として発表する、という形式で運営されているという（吉永 2019）。
- (11) 前掲注3参照。ルーブリックの重要な意義のひとつは教育的効果であり、趣旨は水準で基準を客観的かつ明確にして学生と共有することで、学生の多数が優劣なく満点評価を得ることが理想とされる。
- (12) 要求に対する努力を最小限にしようとする傾向のことで、オンラインの調査においては、教示に適切に従わないなどの行動が挙げられ、その結果、データにバイアスが生じる。したがって、データ測定において努力の最小限化を検出し分析データから除外する必要がある（三浦・小林 2015）。
- (13) 以下、本稿における統計分析の基本的な意味については、たとえば森田 2014 参照。用語法等は可能な限り同書の表現に統一した。
- (14) 大会翌日（11月27日）までに報告されたデータ（94件）のうち、トラップ項目によって努力の最小限化が検出されたデータ（34件）を除外したデータ。なお、2020年度大会では発表は全てアーカイブ化して後日に公開し、学生によるその評価データも大会時と同様に収集しているが、後日評価は分析対象としていない（28日までに報告された評価が対象）。
- (15) ただし5件の中に最優秀賞となった発表は含まれておらず、明らかに低得点となるような発表に対しても高得点がつけられるケースが多くみられた。
- (16) 級内相関係数（ICC）は3以上のデータ間の一致度を示すものであり、0から1までの値をとり、0ならば全く一致しない、1ならば完全に一致することを示す。これを指標にルーブリックの妥当性・信頼性について検証した先行研究としては、貝谷ほか 2017 や、本研究と同様に3名の評価者による検証を詳細に行った例として會田ほか 2018 がある。一般に「評価が一致している」と解釈可能な値は $ICC = 0.7$ 程度以上と解されており（対馬 2010、會田ほか 2018）、貝谷ほか 2017 では $ICC = 0.53 \sim 0.81$ 、會田ほか 2018 では $ICC = 0.792$ との結果であったことから、ルーブリックの妥当性・信頼性が検証されたと結論付けられている。
- (17) 本研究におけるICCは、得点の全体的なばらつき（発表間の分散+評価者間の分散）のうち、発表間のばらつき（発表間の分散）がどの程度の割合を占

研究ノート

めるのかを示したものである（発表間分散 / (発表間分散 + 評価者間分散)）。すなわち、発表間での得点のばらつきが大きく評価者間での得点のばらつきが小さければICCは大きくなり（分子が大きく分母が小さくなり）、反対に、発表間での得点のばらつきが小さく評価者間の得点のばらつきが大きいとICCは小さくなる（分子が小さく分母が大きくなる）。今回、多くの発表が「1」だった、すなわち発表間の得点のばらつきが小さく、一方で、「0」の評価者間での完全一致がない、すなわち評価者間のばらつきが（完全一致がある場合に比べて）大きかったため、ICCが低くなったものと考えられる。

- (18) 2つの変数の関連性を示す指標を「相関係数」といい、「r」と表記される。相関係数は-1.0から+1.0までの値をとり、絶対値が1.0に近づくほど強く関連していることを示す。2つの変数同士に関連が見られないケースでは0を、いずれか一方の得点が高いほどもう一方が高いというケースでは正の値を（正の相関関係という）、いずれか一方が高いほどもう一方が低いというケースでは負の値をとる（負の相関関係）。また、相関係数の有意性検定（少なくとも相関が0ではないことの検証）では、「相関係数が0である」という帰無仮説のもとで実際の相関係数の値が生じる確率を算出し、その確率が5%未満であった場合、帰無仮説を棄却し「相関係数は0ではない（=少なくとも関連性があるといえる）」を採択する。この時の確率を「p」と表記する。なお、教員による評価では、教員1人につき5組の発表を評価しており、各発表は3名の教員によって評価されている。したがって、各評価は教員によってネストされているとともに、発表によってもネストされていることから、教員と発表の変量効果を仮定し、ルーブリック評価を説明変数、総合点評価を目的変数としたマルチレベルモデルを検証した。分析の結果、相関分析の結果と類似していた ($b_{\text{fixed}} = 0.15$, $p < 0.01$ 、分散 $\text{教員} \cdot \text{切片} = 1.74$ 、分散 $\text{教員} \cdot \text{傾き} = 0.003$ 、分散 $\text{発表} \cdot \text{切片} = 1.04$ 、分散 $\text{発表} \cdot \text{傾き} = 0.004$)。
- (19) 学生全評価については、表8・図2に関連して述べた疑念のため、発表ごとの級内相関係数は算出していない。