

*The Journal of
Nagasaki University of Foreign Studies
No.25 2021*

形成的評価としてのルーブリック評価に関するケーススタディ
—留学生による自己評価と他者評価に注目して—

松本 剛次

A Case Study on Rubric Evaluations as Formative Assessments
Focusing on self-assessment and peer-assessment by international students

MATSUMOTO Koji

長崎外大論叢

第25号
(別冊)

長崎外国語大学
2021年12月

形成的評価としてのルーブリック評価に関するケーススタディ —留学生による自己評価と他者評価に注目して—

松本 剛次

A Case Study on Rubric Evaluations as Formative Assessments Focusing on self-assessment and peer-assessment by international students

MATSUMOTO Koji

Abstract

This study discusses hypotheses from the results of self- and peer-assessment of presentations and speeches made by international students using rubric evaluation charts. Furthermore, it attempts to discuss the usage of rubric evaluation in classes. Analyzing data from the Japanese language classes held in the spring semester in the 2021 academic year revealed the following: 1) Numerous students had self-evaluation abilities from the beginning, but naturally, some did not; 2) Both students with low and high self-evaluation abilities adjust their self-evaluation through other students' evaluations; and 3) For international students, the ability to understand the Japanese language is considered to significantly affect their ability to evaluate others. These results confirm the usefulness of rubric evaluations in self-evaluation. The importance of improving self-evaluation ability has also been discussed.

キーワード

学士力 ルーブリック評価 形成的評価

1. はじめに

1.1 目的と意義

近年、テストなどの形で測定可能な、いわゆる「学力」以外のより総合的な「能力」を評価し、また、それをさらに伸ばしていく目的で、ルーブリック評価というものが大きく注目を集めている。特に、2008年に中央教育審議会が「学士力」というものを大学学士課程で身に着けるべき能力の一つの目安として提示して以来、ルーブリックを用いての諸能力の評価は、多くの大学で行われるようになった。留学生も当然例外ではなく、正規留学生はもちろん、本国の大学に籍を置くいわゆる交換留学生であっても、日本国内の大学で学んでいる以上、日本語力の向上のみならず、学士力に代表される、より総合的な能力の向上こそが、大学での教育の目標である。そして、その目標が実現されているかを確認するためにはそれらの「能力」を測定し、評価するためのツールが必要である。その一つがルーブリックである。

しかし、多くの大学において、ルーブリックというものが開発され、使用されるようになってきてはいるものの、多くの場合、それは自己評価ツールとして、学期ごと、あるいは年度ごとに学生自身

の手により、現在の自分の状態を記述しておくものとして使われているが現状であろう¹。もちろんそのような使用にも意味があるし意義もある。しかし、そのように自己評価シートとしてルーブリックを使用するという事は、そもそも学生が十分な自己評価、自己分析ができる、そのような自己評価能力があるということを前提としている。ここに大きな矛盾があることがまずは指摘できよう。客観的な自己評価ができる、自己分析ができるということ自体がひとつの「能力」であり、それこそが大学で育成すべき「能力」の一つである。それなのに、そのような「能力」が既にあるという前提でルーブリック評価というものが進められてしまっているという矛盾である。ルーブリック評価を行うのであれば、まずは評価を行う大学側が、学生側にはそのような自己評価能力があるのかどうかを判断し、ないのであれば、それを何らかの形で育成してやる必要がある。

本研究はこのような問題意識から、まずは学生自身のルーブリックによる自己評価の現状について、それを教師やクラスメートなどの他者からの評価と比較することで見ていくものである。自己評価と他者評価の間に大きな違いがないのであれば、それは、即ち自己による主観的な評価と他者による客観的な評価には差がないということになり、評価結果自体の信頼性（「測定しようとしているものをどれくらい的確に測定できているか」という意味でこの場合は正確には「妥当性」といった方が正しい）はともかく、評価者間での評価が一致するという意味での信頼性（=評定者間信頼性）は高いことを意味する。反対に、自己評価と他者による評価の差が大きい場合は、その信頼性は低いということになる。もちろんその場合、自己と他者のどちらの評価が正しいか（信頼できるか）は本調査の結果のみからは判断できない。よって本調査で見ることが可能なのは、あくまでルーブリックによる自己評価と他者評価の現状、そこに大きな差があるかないかについてのみである。しかし、その結果からを踏まえた上で何らかの仮説を見出し、その仮説から、では今後、それをどう使うのがよいか、というその使い方について考察、提言することは可能である。本調査が目的とするのは、ルーブリック評価自体の「妥当性」や「信頼性」の検証や検討ではない。ましてやそれを踏まえた上での「妥当性」や「信頼性」の高いルーブリックの作成ではない。本調査が目指すのは、本調査の結果を踏まえた上で、ルーブリック評価の使い方について検討することである。

1.2 対象について

本調査の直接の対象は留学生である。留学生を対象としたのは、先にも述べたように、留学生にも広い意味での能力の向上が大学教育においては求められているが、未だに留学生の場合は狭義の意味での「日本語力」の方が評価の対象とされることが多く、広い意味での能力の向上については議論されることが少ないという現状（松本, 2021）を踏まえてのことである。しかし、「学士力」に代表されるような一連の「能力」やルーブリック評価自体については留学生と日本人学生の間で大きな差があるとは考えられない。よって、本調査は留学生を対象としたものではあるが、その結果やその結果から考えられることは日本人学生に対しても応用可能、適用可能なものである。

2. 先行研究－大学におけるルーブリック評価を巡る近年の動向

ここでレビューするのはあくまで「学士力」という考え方が提示されてからの日本国内の大学におけるルーブリック評価に関わる研究である。ルーブリックという考え方、及びその調査方法については詳しくはスティーブンス&レビ（2014）などを参照されたい。

2012年に中教審が『学士課程教育の構築に向けて』を発表して以来、いわゆる「学士力」に代表されるような、それまでの「知識」の獲得からより広い意味での「能力」の獲得へと大学教育の目的がシフトしてきたのは先述のとおりである。そしてそのようなシフトに伴い、当然「評価」の方法やそのありかたもシフトしていくことの必要性も強く唱えられるようになった。

そこで注目されるようになったのがルーブリック評価である。沖(2014)、田宮(2014)は、シフト後比較的早い段階でルーブリック評価の必要性と重要性を唱えていた論文である。沖(2014)ではアメリカの高等教育機関での例としてAssociation of American Colleges & University (AAC&U)が提供しているVALUE Rubricsが紹介されている。一方、田宮(2014)ではStevens & Levi(2005)を参考にルーブリックの意義とその使い方が紹介されている。

「紹介」の次に来たのはその「検討」であり「応用」であった。例えば、山田・森・毛利・岩崎・田中(2015)ではルーブリックの評価に関する方法論の検討が行われた。また、深井・河合・中野(2017)は実際にルーブリック評価を初年次情報教育に取り入れ、その自己省察支援としての有効性を検討した。さらに渡辺・新井・佐々木・盛・古川・水谷・眞坂・塩野目・高井・有本(2019)では先に述べたAAC&UのVALUE Rubricsと河合塾が開発したジェネリックスキル(汎用的能力)の成長を支援するプログラムであるPROGを参考としながら「必要となる汎用的能力の評価指標を明確にし、それらの要件を満たすルーブリックとチェックリスト」を作成し提案している。

そして現在では、アメリカでの先行例などを参考に、各大学が、各大学の理念や教育目標に合わせたルーブリックを作成し、まだ、試験的な意味合いは強いものの、それを使用した評価に取り組み始めている。近年ではその評価の方針(アセスメント・ポリシー)に「学士力に関するルーブリック評価」というものを明記し、各大学が独自に作成したルーブリックを公開し、学生に対する学習成果の評価を実施している大学やコース(プログラム)も、多く見られるようになってきた²。例えば金沢大学の「[人類の未来をつくる独創的研究者の育成]プログラム」(参考資料)は、「現在の最先端科学や最新科学技術はますます複雑化・高度化しており、細分化された学問領域一つの知識や専門スキルを学ぶだけでは、課題を解決することが極めて困難になってきている」との現状認識のもと、「従来の発想にとらわれず、現象の背後に潜む法則性を見つけ、分野融合的な視座から解決の糸口を探る力」「助言や批判を受け止めながらも、自己の根幹の考えにこだわり真理に迫る力」「世界で活躍するための力」することが目指されている高校生を対象とした高大接続プログラムであるが、そこでは独自のユニークなルーブリックが作成され公開されている³。

また、東北福祉大学では学士力関連コモングルブリック、科目ごとのルーブリック、授業コモングルブリックという一連のルーブリックが作られ、「Web版の学習ポートフォリオ」において学生の自己評価による「振り返り」としてルーブリック評価を用い、その平均値を学年推移、及び入学時と卒業時を比較することでその評価が行われている(参考資料)。なお、学士力関連コモングルブリックには「コミュニケーション能力(ディスカッション能力を含む)」「プレゼンテーション能力」「問題発見・解決能力」「情報リテラシー」「態度・志向性」「レポート力」「数量的リテラシー」「人類の文化・社会と自然に関する知識の理解」「多文化・異文化に関する知識の理解」「専門性(専門的な知識・技能や技術・倫理)」の10の下位分類があり、それぞれにルーブリックが作成されている。

3. データ

3.1 本調査の位置づけと使用したルーブリックについて

第2章で述べた進展を踏まえた上で、第1章で述べた問題意識の下行われるのが、本調査である。「紹介」「検討」「応用」という段階で進展してきたルーブリック評価を巡る議論を、次の段階の「使用」へと進めるための研究に向けてのパイロットスタディであるとも言えよう。

しかし既に述べたことの繰り返しにもなるが、本研究は、少なくとも現時点においては、ルーブリック自体の作成を目的としたものでもないし、その（作成されたルーブリックの）信頼性や妥当性について調査するものでもない。本研究はあくまでルーブリックによる自己評価と他者評価の現状についての調査であり、そしてその結果から考察されるのは、その現状を踏まえた上で、今後、それをどう使うのがよいか、というルーブリックの使い方のほうである。そこで、本研究では、既に公開されている既存のルーブリックを使用した。具体的には、先に紹介した東北福祉大学の一連のルーブリックの中から、「コミュニケーション能力（ディスカッション能力を含む）」「プレゼンテーション能力」「問題発見・解決能力」の3つのルーブリックを使用した。この3つを選んだのは、本研究の舞台となった授業が「総合日本語」という授業で、そこでは実際にディスカッションやプレゼンテーションという活動が多く行われたからである。また、「問題発見・解決能力」は数多くの大学生に必要とされる諸能力の中でも、その中核を担うものと考えられるからである。

次の表1-表3が本研究において使用したルーブリックの一覧である。なお、当該ルーブリックの使用にあたり、東北福祉大学にはメールで連絡し、使用についての許可をもらっている。

3.2 対象と調査が行われた授業、及びデータの収集方法について

本研究のデータが採取された授業は、2021年4月から8月までの2021年度春学期に行われた「総合日本語」の2つのレベルのクラスである。総合日本語は読む、聴く、話す、書く、の4技能のうち、特に「話す」ことに主眼を置いたクラスである。具体的には毎回テーマを決めた上で（同じテーマが複数回にわたり続く場合もある）グループ及びクラス全体でのディスカッションやそのディスカッションを踏まえた上での個人の意見を発表（プレゼンテーション）し合う、という活動が繰り返し行われた。また、ビブリオバトルスタイルで、自分が読んだ本や見た映画、ドラマ、やってみたゲーム

表1：東北福祉大学での「問題発見・解決能力」のルーブリック

到達目標	評価する能力	内容	5	4	3	2	1
1 物事のあるべき姿と現状のギャップについて把握できており、何が解決すべき問題なのかを正確に見つけ出している。	課題発見力	課題について	課題に対してのあるべき姿と現状のギャップについて把握でき、何が解決すべき課題なのかを見つけ出すことができる。	課題を正しく理解し、物事のあるべき姿と現状のギャップについて把握できる。	課題発見解決というテーマを理解し、物事のあるべき姿と現状のギャップについて把握しようとしている。	課題発見解決力というテーマを理解できる。	課題発見解決力というテーマを理解できていない。
2 調査・分析によって何が課題かを因果関係から理解し、解決のための方向付けと優先順位が付けられている。	課題分析力	調査・分析について	調査・分析によって何が課題かを因果関係から理解でき、解決のための方向付けと優先順位を付けることができる。	調査・分析によって何が課題かを因果関係から理解でき、解決のための方向付けと優先順位をある程度付けられる。	調査・分析によって何が課題かを因果関係から理解でき、解決のための方向付けができる。	調査・分析によって何が課題かを因果関係から理解できる。	調査(情報収集)・分析が不十分である。課題の因果関係を理解できない。
3 達成目標を設定し、順序立てて計画を立て、実行している。	計画実行力	計画・実行について	達成目標を設定し、創造的思考をもって最善の計画を立て、その計画に沿って実行できる。	達成目標を設定した計画を立てることができる。	達成目標を設定し、計画を立てて実行できる。	計画を立てることができるが、達成目標が曖昧である。または具体的なは無理がある。あるいは、計画倒れで実行が伴わない。	計画を立てることができない。
4 実行した結果を多面的に検証して正しく評価し、必要ならば改善し、次の計画へと反映させている。	検証力	検証について	実行した結果を多面的に検証し、正しく評価し、必要ならば改善し、次の計画へと反映させることができる。	実行した結果を検証し、正しく評価し、次の計画へとつなげることができる。	実行した結果を検証し、正しく評価できる。	一通り検証ができる。	検証をしていない。

について紹介し合う、という活動も行われた。便宜的にここでは、その2クラスをAクラス、Bクラスとして紹介する。Aクラスは全員で3名、全てが中国出身の留学生という構成であった。Bクラスは全員で7名、内訳は韓国出身の留学生が2名、中国出身の留学生が5名という構成であった。

表2：東北福祉大学での「コミュニケーション能力(ディスカッション能力を含む)」のルーブリック

到達目標	評価する能力	下位項目	5	4	3	2	1
A (少人数ゼミ、実習でのディスカッション、ディベート等においてグループの中で他人の意見を尊重して聞き、正しく理解し、かつ批判的に考えたりして、自分の意見を分かりやすく表現できる。	傾聴力、表現力、相互理解	1 他人の意見について	尊重して聞き、十分に理解した上で、さまざまな視点から考えることができる。	尊重して聞き、深いところ(意図や気持ちなど)まで理解できる。	尊重して聞き、言いたいことを正しく理解できる。	尊重して聞き、一通り理解できる。	尊重していないか、鵜呑みにしており、理解できていない。
		2 自分の意見について	さまざまな説明の方法や手段を駆使し、わかりやすく伝え、意見の異なる相手との相互理解を得ることができる。	わかりやすく伝え、意見の異なる相手との相互理解を得ることができる。	伝えることができるが、相互理解には到っていない。	それなりに伝えることができる。	伝えていないか、伝わっていない。
B (少人数ゼミ、実習でのディスカッション、ディベート等において専門的なテーマに関して十分な学習をし、グループでの討論や議論において適切な手順・手段を用いて分かりやすく説明したりして、グループの中で意見を自分なりにまとめ、今までになかった新しいアイデアの発想へとつなげることができる。	情報収集力、要約力、創造力	1 専門的なテーマに関する学習について	さまざまな角度から十分に学習を行っている。	十分に行っている。	適度に行っている。	行っているが、不十分である。	行っていない。
		2 議論の目標について	設定し、それに合わせてさまざまな角度から議論を展開できる。	設定し、それに合わせて議論を十分に展開できる。	設定し、それに合わせて議論を適度に展開できる。	設定し、それに合わせた展開を心がけているが、不十分である。	設定せず、一方的な主張・誤った認識によって議論になっていない。
		3 グループでの討論や議論における自分なりの意見や主張について	適切で創造的な手順・手段を用いて分かりやすく説明し、相手を十分に納得させることができる。	適切な手順・手段を用いて分かりやすく説明し、相手を納得させることができる。	適切な手順・手段を用いて分かりやすく説明できる。	適切な手順・手段を用いて説明しようとしているが、不十分である。	適切な手順・手段を用いて分かりやすく説明していない。
		4 グループの中で意見について	自分なりにまとめ、グループ全体を今までになかった新しいアイデアの発想へとつなげることができる。	自分なりにまとめ、要点を正確に述べることができる。	自分なりにまとめ、要点を正確に述べることができる。	自分なりにまとめているが、不十分である。	自分なりにまとめている。

この2クラスにおいて、「問題発見・解決能力」と「コミュニケーション能力(ディスカッション能力を含む)」については、第2回目の授業と中間段階である第8回目の授業、及び最終回である第15回目の授業において学生自身の手による当該ルーブリックを用いての自己評価が行われた⁴。第1回目ではなく第2回目なのは、第1回目はオリエンテーションという位置づけであり、まだ受講者が定まっていなかったからである⁵。そしてこの「問題発見・解決能力」と「コミュニケーション能力(ディスカッション能力を含む)」について、学期終了の時点で教師である筆者によるそれぞれの学生に対

表3：東北福祉大学での「プレゼンテーション能力」のルーブリック

到達目標	評価する能力		5	4	3	2	1
A 1) 課題の内容に対する歴史的な背景や周囲の状況を把握でき、課題の目的が明快に説明できる。 2) 調査・研究方法・結果をその道筋に沿って説明できる。 3) 十分な検証のもと、結果を課題と関連づけて説明し、結果の持つ意味合いを理解している。 4) 課題解決法の特徴・オリジナリティを示し、今後の展望について述べることができる。	発表事前準備力	1 課題の内容に対して	十分に歴史的な背景や周囲の状況を調べ、さまざまな角度から考えて把握でき、課題の目的を明快に説明できる。	歴史的な背景や周囲の状況が正確に把握できており、課題の目的を明快に説明できる。	歴史的な背景や周囲の状況が把握でき、課題の目的を説明できる。	学習しているが、不十分である。	学習が足りていない。
		2 調査・研究方法・結果について	その道筋に沿って効果的に説明でき、優れたオリジナリティを示すことができる。	その道筋に沿って効果的に説明できる。	その道筋に沿って説明できる。	説明しているが、不十分である。	説明していない。
		3 結果と課題の関連づけについて	関連づけて説明し、結果の持つ意味合いを十分に検証したうえで理解できる。	関連づけて説明し、結果の持つ意味合いを検証し理解できる。	関連づけて説明し、結果の持つ意味合いを理解できる。	ある程度関連づけることができる。	関連づけていない。
		4 課題解決法と今後の展望について	課題解決法の特徴・オリジナリティを示し、今後の展望について説得力をもって述べるることができる。	課題解決法の特徴・オリジナリティを示し、今後の展望について述べるることができる。	課題解決法を示し、今後の展望について述べるることができる。	述べているが、不十分である。	述べていない。
B 1) 資料やOHP等が適切に用意できる。 2) 発表の道筋をよく理解しており、堂々とした態度で発表できる。 3) 質問の意味と意図を正確に把握して的確な答えをスムーズに話すことができる。	発表表現力	1 資料やスライド・OHP等について	相手に理解させようと努力してオリジナリティがあるものを適切に用意している。	相手に理解させようと努力して適切に用意している。	適度に用意している。	用意しているが、不十分である。	用意していない。
		2 発表の道筋と発表の態度について	発表の道筋をよく理解しており、堂々と発表し(声、表情を含む)、相手から十分な理解を得ることができる。	発表の道筋をよく理解しており、堂々と発表する。	発表の道筋をよく理解しており、問題なく発表する。	発表の道筋を理解しており、問題なく発表する。	発表の道筋がたつておらず、発表の態度に問題がある。
		3 質問について	質問の意味を正確に把握して的確な答えをスムーズに話すことができる。	質問の意味を正確に把握して的確な答えを話すことができる。	質問の意味を正確に把握して答えを話すことができる。	質問の意味を把握して答えているが、不十分である。	質問の意味を把握していない。的確な答えをしていない。

する他者評価も行われた。なお、学生本人による自己評価の結果が影響しないよう、担当教師である筆者は、各学生が行った自己評価については本論文執筆時まで一切見ていない。

一方、「プレゼンテーション能力」については、まず、「問題発見・解決能力」と「コミュニケーション能力（ディスカッション能力を含む）」の場合と同様に第2回目、第8回目、第15回目の授業で、ルーブリックを用いての自己評価が行われた⁶。その上でさらに、学生が発表（プレゼンテーション）を行った際に、クラスメート同士、及び教師である筆者による発表者への匿名での他者評価も行われた。このような他者評価の活動が行われたのは第4回、第13回、の2回である。他者評価は、授業中に発表者に全員からのものを確認してもらった上で回収し、コピーを取った上で、各自振り返りができるよう、翌週返却した。

なお、本研究を行うにあたり、被験者となる受講生には研究の目的を説明した上で、実名を明らかにしない、個人が特定できるようなことはしない、研究への参加・不参加は自由でありそれが成績等に影響することはないし、後に辞退することもできる、という旨をまずは口頭で十分に説明した。そしてその後、その旨が書かれた同意書を確認してもらい、同意するものには署名をもらった上で回収した。結果としてBクラスの1名を除く全9名から同意書を得ることができた。よって、この9名に担当教師である筆者を加えた10名が本研究における被験者（データ提供者）である。

4. 結果と考察

4.1 「プレゼンテーション能力」のルーブリックを用いた自己評価と他者評価

4.1.1 縦断的分析の結果から

以下、まずは、他者評価も行いデータ量の多い、「プレゼンテーション能力」から結果を見ていきたい。次の表4はAクラス3名、Bクラス6名の自己評価と他者評価の平均を実施順に左から右へと並べていったものである。なお、使用したルーブリックは5段階評価であるが、その中間、たとえば3と4の間に丸が付けられている場合もあり、それは3.5という点数として集計した。

この結果から、まずは自己評価に注目すると、ほとんどの被験者がプラスマイナス1の範囲内で自己評価の調整を繰り返していることが確認出来る。そして他者評価とのずれもそのほとんどがプラスマイナス1の範囲内である。他者評価の信頼性については次項で改めて検討するが、表4を見る限りは、まずは、多くの者がある程度客観的な自己評価を最初の段階から行うことができているとは言えよう。

しかしここで敢えて注目したいのはその例外的なケースであるA02、及びB01である。両者ともに第1回目の自己評価と第2回目の自己評価において2ポイント以上の差が確認できる。A02を見てみると、第2回目の授業で行った初の自己評価においては、「課題解決法と今後の展望について」以外は、2-3と低い自己評価であった。しかし4回目の授業で行われた他者評価の結果を受けて（他者評価は一部を除いて本人の自己評価より高いものであった）、第8回目に行った自己評価ではほとんどが3-4に引き上げられている。また、一方で、第1回目時点での自己評価の高かった「課題解決法と今後の展望について」は引き下げられている。

同様の動き、他者評価の結果を踏まえての自己評価の大幅な修正という動きは、B01にも確認できる。この2名はその差が大きかったからはっきりとその動きが確認できたが、しかし、程度の差こそあれ、他者評価の結果を受けて自己評価を修正するという動きは他の被験者についても指摘できる。ここか

表4：「プレゼンテーション能力」についての全被験者の自己評価と他者評価の縦断的推移

	大分類	下分類	自己	他者平均	自己	他者平均	自己
			2回目	4回目	8回目	13回目	15回目
A01	発表事前準備力	課題の内容に対して	4	3.50	4	4.00	3.5
		調査・研究方法・結果について	4	3.00	4	4.00	4
		結果と課題の関連付けについて	3	2.67	3	3.67	3
		課題解決法と今後の展望について	3	3.00	3	3.50	3
	発表表現力	資料やスライド・OHP等について	4	3.33	4	4.17	4
		発表の道筋と発表の態度について	3	3.33	4	3.67	4
質問について		3	3.50	4	4.00	3	
A02	発表事前準備力	課題の内容に対して	2	4.50	4	4.00	4
		調査・研究方法・結果について	3	4.33	4	4.00	4
		結果と課題の関連付けについて	3	4.33	3	4.00	3
		課題解決法と今後の展望について	5	4.50	3	3.83	5
	発表表現力	資料やスライド・OHP等について	3	4.17	4	4.00	4
		発表の道筋と発表の態度について	2	4.17	3	3.50	4
質問について		2	3.67	3	4.00	4	
A03	発表事前準備力	課題の内容に対して	3	3.67	4	4.33	4
		調査・研究方法・結果について	3	4.00	4	4.33	4
		結果と課題の関連付けについて	3	4.00	3	4.00	4
		課題解決法と今後の展望について	3	4.00	3	3.83	4
	発表表現力	資料やスライド・OHP等について	4	4.00	4	4.00	3
		発表の道筋と発表の態度について	3	4.00	4	3.50	3
質問について		3	3.67	2	4.33	3	
B01	発表事前準備力	課題の内容に対して	3	4.40	5	4.43	5
		調査・研究方法・結果について	3	4.60	4	4.57	5
		結果と課題の関連付けについて	3	4.40	4	4.57	5
		課題解決法と今後の展望について	4	4.20	4	3.86	5
	発表表現力	資料やスライド・OHP等について	3	4.60	5	4.00	4
		発表の道筋と発表の態度について	4	4.50	5	4.43	5
質問について		4	4.40	5	4.43	5	
B02	発表事前準備力	課題の内容に対して	4	4.40	4	4.57	4
		調査・研究方法・結果について	4	4.00	4	4.43	4
		結果と課題の関連付けについて	3	4.50	4	4.43	4
		課題解決法と今後の展望について	4	4.00	3	4.43	4
	発表表現力	資料やスライド・OHP等について	3	4.30	3	4.71	4
		発表の道筋と発表の態度について	3	4.60	3	4.57	4
質問について		4	4.20	4	4.57	4	

B03	発表事前準備力	課題の内容に対して	3	4.42	3	5.00	5
		調査・研究方法・結果について	4	4.42	4	5.00	5
		結果と課題の関連付けについて	4	4.58	4	4.86	5
		課題解決法と今後の展望について	3	4.42	4	4.86	5
	発表表現力	資料やスライド・OHP等について	5	4.17	5	4.71	4
		発表の道筋と発表の態度について	5	4.67	4	5.00	5
		質問について	4	4.75	4	5.00	4
B04	発表事前準備力	課題の内容に対して	4	4.67	5	4.71	5
		調査・研究方法・結果について	3	4.50	4	4.57	4
		結果と課題の関連付けについて	4	4.58	5	4.57	5
		課題解決法と今後の展望について	4	4.67	4	4.57	5
	発表表現力	資料やスライド・OHP等について	4	4.83	5	4.71	5
		発表の道筋と発表の態度について	3	4.58	5	4.43	4
		質問について	4	4.67	4	4.50	5
B05	発表事前準備力	課題の内容に対して	3	4.00	3	4.17	3
		調査・研究方法・結果について	4	3.80	3	4.00	3
		結果と課題の関連付けについて	3	4.10	3	3.83	3
		課題解決法と今後の展望について	3	3.90	3	4.00	3
	発表表現力	資料やスライド・OHP等について	3	3.70	3	3.83	3
		発表の道筋と発表の態度について	3	4.10	3	3.83	3
		質問について	3	3.75	3	4.25	3
B06	発表事前準備力	課題の内容に対して	3	—	4	4.67	4
		調査・研究方法・結果について	3	—	4	4.83	4
		結果と課題の関連付けについて	4	—	5	4.67	4
		課題解決法と今後の展望について	3	—	5	4.50	4
	発表表現力	資料やスライド・OHP等について	4	—	5	4.67	4
		発表の道筋と発表の態度について	3	—	5	4.67	5
		質問について	4	—	5	4.83	5

ら、基本的には多くの者が最初の段階から客観的（他者の視点からによる他者評価に近いという意味での「客観的」）な自己評価というものを行うことができているが、中にはそれが苦手な者も当然ながらいることもまた事実であることがまずは確認できる。そして同時に、その自己評価が苦手な者も、他者評価によるフィードバックを受けることによって、自らの評価（＝自己評価）を調整することができるようになる、という状況もここから見て取ることができる。そして、この、「他者評価によるフィードバックを受けることによって、自己評価を調整することができるようになる」という考え方からは、「他者評価を受けるだけでなく、それ（他者評価）を行うということも、自らの位置というものも客観的に捉える能力の育成に寄与するのではないか」という仮説も引き起こされる。事実、A02についても、B01についても、他者評価を受けたから自己評価が変わったのか、他者評価を行ったから、自己評価も変わったのか、それは、今回のデータからだけでは何とも言うことはできない。言えるのは他

者評価を受けた/行った経験をした後で、自己評価も変化したという事実だけである。

なお、それでもさらに例外的なものとしてB05とB06の事例も検討してみる必要があるだろう。B05は他者評価の結果を受けても自己評価は常にすべてが3のままである。私見ではあるが、このB05は出席率も低く、授業態度も良いとは言えない（人の発表をきちんと聞いているとは思えない）タイプの学生であった⁷。他者評価を返却しても、それにしっかりと目を通していなかったのではないかと思われる。一方、B06は4回目の授業を欠席したために、他者評価というフィードバックが得られていないし、自身も他者への評価をその時は行っていない。それでも2回目の自己評価から8回目の自己評価の間で、かなりの自己評価の引き上げが行われている。そしてそれは13回目に行われた他者評価の結果に近いものとなっている。これも私見であり、想像の範囲を出ないものであるが、B06はB05とは対照的ないわゆる真面目な学生であり、他者の発表にも熱心に耳を傾けるタイプの学生であった。自己評価、他者評価を行った回以外にも、様々な形で学生による発表の類は一連の授業の中で行われていた。先に、「他者評価によるフィードバックを受けることによって、自己評価を調整することができるようになる」、「他者評価を受けるだけでなく、それ（他者評価）を行うということも、自らの位置というものも客観的に捉える能力の育成に寄与するのではないか」という仮説を提示したが、実際にループリックを使用しての他者評価を行わなくても、他者の発表を積極的に聞くだけでも、同様の効果（自分の現在の能力がどの程度か、今現在どの位置にあるのか、ということ客観的に把握できるようになること）が期待できるのではないだろうか。その可能性をB06の事例は示唆しているようにも思われる。

4.1.2 他者評価の信頼度の観点から

以上まとめると、表4のデータからは、「当初から客観的な自己評価というものができものがいる一方でそうではない者も当然存在し、そのような者たちは、自己評価と他者評価を繰り返し行うことで客観的な自分の位置づけ（=自己評価）というものを次第に把握できるようになっていくのではないか」という仮説が考察された。他者評価については、それを「受ける」とそれ「行う」ととの間で、どちらが客観的な自己評価の能力を育成することにより効果的なのかについては、このデータのみからは何とも言えない。しかし、では、その他者評価はどの程度信頼できるものだろうか。それ（他者評価）が信頼できないものなのであれば、その結果得られる自己評価の能力も信頼できるものではなくなってしまう。既に何度も繰り返し述べているように、本調査はループリック自体の信頼性を調査するものではない。しかし、複数のものが一人に対して他者評価を行った場合、そのそれぞれが行う他者評価の間にばらつきが少ないのであれば、少なくとも他者に対する評価は、多くの者が同じように行うことができるということになり、それは即ち、他者評価はある程度の信頼性（評定者間信頼性）を有している、ということにはなる。

ここでは、その「ばらつき」を各評価者が付けた評価点の標準偏差（SD）を見ることによって確認したい。次の表5がそれである。

ここでまず注目したいのは、標準偏差が1.0を超える例はAクラスにのみ認められるという事実である。しかし、これは一種の天井効果であるとも言える。5段階評価であることから、レベルが上のBクラスは評価が4か5に集中してしまい、3や2といった評価もあり得るAクラスに比べると、そもそも差が開かないからである。しかし、一方ではレベルが下のAクラスであっても、集中する範囲と

表5：各被験者への他者評価の標準偏差

		発表事前準備力				発表表現力		
		課題の内容に対して	調査・研究方法・結果について	結果と課題の関連付けについて	課題解決法と今後の展望について	資料やスライド・OHP等について	発表の道筋と発表の態度について	質問について
対A01	他者評価①4回目のSD	0.87	0.00	0.58	0.00	0.58	0.58	0.50
	他者評価②13回目のSD	1.00	0.00	0.58	0.50	0.29	0.58	0.00
対A02	他者評価①4回目のSD	0.50	0.58	1.15	0.50	0.29	0.29	0.58
	他者評価②13回目のSD	0.00	0.00	0.00	0.76	1.00	0.50	0.00
対A03	他者評価①4回目のSD	0.58	0.00	1.00	0.00	0.00	0.00	0.58
	他者評価②13回目のSD	0.58	0.58	1.00	0.29	0.00	0.50	0.58
対B01	他者評価①4回目のSD	0.55	0.55	0.55	0.45	0.55	0.50	0.55
	他者評価②13回目のSD	0.53	0.53	0.53	0.69	0.82	0.53	0.53
対B02	他者評価①4回目のSD	0.55	0.71	0.50	0.71	0.84	0.55	0.45
	他者評価②13回目のSD	0.53	0.53	0.53	0.53	0.49	0.53	0.53
対B03	他者評価①4回目のSD	0.80	0.49	0.49	0.80	0.75	0.52	0.42
	他者評価②13回目のSD	0.00	0.00	0.38	0.38	0.49	0.00	0.00
対B04	他者評価①4回目のSD	0.41	0.45	0.49	0.52	0.41	0.49	0.52
	他者評価②13回目のSD	0.49	0.53	0.53	0.53	0.49	0.53	0.50
対B05	他者評価①4回目のSD	0.71	0.84	0.55	0.74	0.84	0.89	0.96
	他者評価②13回目のSD	0.75	0.63	0.98	0.89	0.98	0.98	0.50
対B06	他者評価①4回目のSD	—	—	—	—	—	—	—
	他者評価②13回目のSD	0.52	0.41	0.52	0.55	0.52	0.52	0.41

いう意味ではレベルが低いなら低いなりに例えば、3と4などとなり、その意味ではBクラスと変わらないのではないかと、という指摘もあろう。また他者評価を行う人数もAクラスは自分以外の3名（教師である筆者も含む）、Bクラスは6名となり、評価に差が出る可能性が高いのはむしろ評価する人数の多い、Bクラスのほうであろう。

これも私見ではあるが、このような結果となったのにはやはり日本語力、特に日本語での理解力が大きく関係していると考えられる。そもそも発表の内容が十分に理解できていなければ正しい評価を行うことはできない。Aクラスは他者評価の能力が低いということよりも、日本語の理解力が低いことがその評価にばらつきの多い（分散が高い）原因として考えられる。一方、日本語力が相対的に高いBクラスは内容をお互いに十分に理解できるが故に、他者評価も安定している（ばらつき=分散が少ない）。4.1.1において「多くの者がある程度客観的な自己評価を最初の段階から行うことができている」と述べたが、他者評価については、留学生の場合、日本語能力、特に日本語での発表の理解の能力が大きく影響しているのではないかと考えられる。

4.2 「コミュニケーション能力」と「問題発見・解決能力」のルーブリックを用いた自己評価と他者評価

以上、①当初から自己評価能力はある者が多いが、そうではないものも存在すること。②自己評価能力が低い者も高い者も他者評価を受けることで、あるいは他者評価を行うことを通して、自己評価

のほうの調整も行っていること、そしてそれを通して自己評価能力をさらに伸ばしていること、③一方、他者評価能力については、留学生の場合は日本語の理解能力が大きく影響していると考えられること、の3つが、あくまで「仮説」ではあるが、「プレゼンテーション能力」についてのルーブリックを用いた自己評価と他者評価の結果から検討、考察された。では、他者評価によるフィードバックや他者評価を行う機会がなかったものについてはどうだろうか。ここでは「コミュニケーション能力（ディスカッション能力を含む）」のルーブリックを用いた自己評価と同じルーブリックを用いた教師（=筆者）による他者評価、及び、「問題発見・解決能力」のルーブリックを用いた自己評価と他者評価（同じく教師である筆者による）の結果について見てみたい。

紙幅の都合上、これらについては各クラス1名、A01とB01の2名の結果のみを以下に表6、表7として示すに留める。後に述べる1例を除き、そのほとんどがこのようにプラスマイナス1の範囲内で自己評価の調整を繰り返し、最終段階（15回目）に行われた教師である筆者による他者評価との間の差も、プラスマイナス1の範囲内に収まるものであったからである⁸。「コミュニケーション能力（ディスカッション能力を含む）」のルーブリックを用いた評価においても、「問題発見・解決能力」のルーブリックを用いた評価においても、4.1.1で指摘した「多くの者がある程度客観的な自己評価を最初の段階から行うことができている」という事実を改めて確認することができる。

表6：「コミュニケーション能力（ディスカッション能力を含む）」の自己評価と他者評価の例

			1週目	8週目	15週目	教師による他者評価（15週目）
A01	傾聴力、表現力、相互理解	他人の意見について	3	4	4	3.5
		自分の意見について	4	3	3	3.5
	情報収集力、要約力、創造力	専門的なテーマに関する学習について	3	4	4	3.5
		議論の目標について	3	3	3.5	3.5
		グループでの討論や議論における自分なりの意見や主張について	3	3	3	3.5
		グループの中での意見について	3	3	3	3

B01	傾聴力、表現力、相互理解	他人の意見について	4	5	5	4
		自分の意見について	5	5	5	4
	情報収集力、要約力、創造力	専門的なテーマに関する学習について	3	4	4	3.5
		議論の目標について	3	4	4	3.5
		グループでの討論や議論における自分なりの意見や主張について	5	4	5	4
		グループの中での意見について	4	5	5	4

表7：「問題発見・解決能力」の自己評価と他者評価の例

		1週目	8週目	15週目	教師による他者評価（15週目）
A01	問題発見力	3	3	3.5	4
	課題分析力	2	4	3	3.5
	計画実行力	3	3	3.5	4
	検証力	2	3	3	3.5
B01	問題発見力	4	4	5	4
	課題分析力	3	3	4	4
	計画実行力	3	3	5	4
	検証力	3	4	4	3.5

しかし、これも4.1.1で確認したように、自己評価が苦手な者も、少数ながらやはり存在していることもまた事実である。次に表8として示すのはA02の「問題発見・解決能力」の評価結果である。A02は4.1.1でも自己評価が苦手な者の例として挙げた者であるが、ここでも最終的な自己評価をかなり高くしており、教師である筆者との評価（この両者を太字で示す）との間には差が認められる。「コミュニケーション能力（ディスカッション能力を含む）」「問題発見・解決能力」については、他者評価（=教師である筆者による評価）との間に大きな差が認められたのはこの1事例のみであるが、これまでに検討してきたことと併せて考えると、最後の15週目にしてこのような結果になったということは、少なくともこの1名についてはやはり途中で他者評価を受けること、あるいは行うことによるフィードバックが必要だったことを示唆していると考えられる。

表8：自己評価と他者評価の差が大きいものの事例

A02	問題発見力	4	4	5	3.5
	課題分析力	3	4	4	3.5
	計画実行力	4	3	4	3.5
	検証力	2	3	5	3.5

5. まとめと提言

以上、「プレゼンテーション能力」のルーブリックを用いた自己評価と他者評価の結果から①当初から自己評価能力はある者が多いが、そうではないものも存在すること。②自己評価能力が低い者も高い者も他者評価を受けることで、あるいは他者評価を行うことを通して、自己評価のほうの調整も行っていること、そしてそれを通して自己評価能力をさらに伸ばしていること、③一方、他者評価能力については、留学生の場合は日本語の理解能力が大きく影響していると考えられること、の3点が「仮説」として見いだされ、このうち①については「問題発見・解決能力」のルーブリックを用いた自己評価と他者評価の結果からも確認することができた。

1.1で言及していたように、本調査が目指すのは、本調査の結果を踏まえた上でのルーブリック評価の使い方について検討することである。では、この結果から、ルーブリック評価の使い方についてどのようなことが提案できるだろうか。本論のまとめとしてこの点について考えてみたい。

まず、指摘できるのが、自己評価シートとしてルーブリックを使用することは可能であるが、全員が最初の段階から客観的な自己評価ができるわけではないという事実である。そしてその事実から、「だからこそ、自己評価のトレーニングが必要である」、という見解を提案することができる。そしてそのトレーニングの際に役立つのが、本調査の結果から仮説的ではあるが見いだされた「他者評価を相互に行うこと」であり、また、「その際に他者から受けた評価をしっかりと受け止めた上で自己評価と照らし合わせる」という作業となる。つまり、ルーブリック評価というものを、単なる現時点での能力の評価のためのツールとして使うのではなく、評価能力を高めるためのツールとして捉える、という考え方、捉え方である。自分自身、及び他者に対してルーブリックを用いた評価を行い、さらにそこからのフィードバックを通して、それを自分自身の成長へとつなげていくこと、このような評価についての考え方、捉え方は、まさに形成的評価の考え方そのものである。ルーブリック評価をまずは自分自身の自己評価能力を高めるためのツールとして使い（形成①）、そしてそのような自

己評価能力を身に着けた上で、今度はさらにルーブリック評価をその本来の目的である自分をさらに上の段階に引き上げるもの（形成②）として使用する、という考え方、捉え方である。「プレゼンテーション能力」を例に挙げれば、まずは、「プレゼンテーション能力」のルーブリック評価表を自己評価、他者評価の際の参照枠として用いることで、自己評価能力（他者評価能力も含む）を高めた上で（形成①）、次に、そのルーブリック評価表をいわゆる指標として、よりよいプレゼンテーションができるようになるためのツール（物差し）として使う（形成②）という考え方である。なお、形成①については、早ければ早いうちに行う方がその後のことを考えると効果的であり、また効率的であると言える。第2章において金沢大学の「人類の未来をつくる独創的研究者の育成」プログラムのルーブリック評価の事例を紹介したが、これは国大接続プログラムとして高校生を対象に行われているものであった点に改めて注目したい。

しかし、留学生の場合、他者評価能力は日本語能力、特に日本語での理解の能力と関係しているのではないということもまた、本調査の結果見えてきた仮説である。大学での活動の多くが日本語で行われることという現状を考えると、日本語力が不十分なうちは正確な他者評価を行うことも難しく、それ故に他者評価の経験が自己評価能力の向上につながる可能性（形成①の可能性）も低くなってしまう。この問題を補うためには、日本語力の低い留学生には、母語や日本語以外の得意な言語を用いてのプレゼンテーションを行わせ、それを相互評価の対象とするなども一案であろう。幸い当大学は外国語大学であり、英語をはじめ、中国語、韓国語、ドイツ語、フランス語を専修している学生も多い、それらを扱う授業も多い。留学生がそれらの言語の授業に出向いて、何らかの発表を行うことは、そしてそれを聞き、それを評価することは、日本人学生にとっても有意義な体験となるであろう。

第1章でも述べたように「学士力」に代表される総合的な能力を育成するのが今日の大学教育に求められている教育のゴール（目標）である。そして言語能力というものは決してそのような総合的な能力とは独立してそれだけで孤立したものとして存在しているのではなく、まさに人は言語を使って考え、行動するという意味で、言語能力こそがその総合的な能力そのものであると言える。言語の教育こそが総合的な能力の教育の中核となり得るものなのである。そしてそのような総合的な能力とは、経験とそのフィードバック、及びそれに対する振り返り（内省、省察）を行うことにより育成されるものである。ルーブリック評価はまさにそのために、自分自身や他者に対するフィードバック、振り返り、内省、省察のためにこそ使われるべきものである。本調査の結果とそこから導き出された仮説や提案は、改めてこの事実を確認させてくれるものである。

注

- 1 本学ではASM（アセスメンター）システムというものを利用しての自己評価と振り返りが学期ごとに行われている。
- 2 2017（平成29）年度から、すべての大学に対して、「入学者の受入れに関する方針」（アドミッション・ポリシー）、「教育課程の編成及び実施に関する方針」（カリキュラム・ポリシー）、「卒業の認定に関する方針」（ディプロマ・ポリシー）という「三つの方針」の策定と公開が義務付けられた。
- 3 同プログラムは令和2年から開始されたものであるが、その前身は平成28年から行われていた「世界でかがやく科学技術イノベーション人材の育成」プログラムである。その時点からルーブリックは作成され公開されていた。

- 4 ここで「週」ではなく「回」という言い方をしているのは、新型コロナウイルスの影響で休講となり後に補講が行われた授業もあったからである。
- 5 第2回目の授業に欠席した者に対しては第3回目の授業で自己評価が行われた。
- 6 同上
- 7 事実B05は最終課題であるクラス内でのビブリオバトルに参加せず、この授業の単位を落としている。
- 8 もちろんAクラスもBクラスも4.1で見た「プレゼンテーション能力」の他者評価の活動は行って、その効果が波及したこともその原因としては考えられる。しかし、少なくとも第1回目の自己評価についてはそれらの活動を行う前のものであるし、次に述べるA02の事例はそのような波及効果(=転移学習)が行われていない事例でもある。

【参考文献】

- 青木久美子 (2017) 「「新しい」大学教育-コンピテンシーに基づく教育 (CBE) の実践」『日本労働研究雑誌』 687.37-45.
- 沖裕貴 (2014) 「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』 14.71-90.
- ダネル・スティーブンス&アントニア・レビ (2014) 『大学教員のためのルーブリック評価入門』 玉川大学出版部 (Stevens,D.D.&Levi,SIJl(2005). *Introduction to Rubrics: An Assessment Tool to Aave Grading Time, Convey Effective Feedback and Promote Student Learning*. Stylus Publishing, LLC.)
- 田宮憲 (2014) 「ルーブリックの意義とその導入・活用」『高等教育開発センターフォーラム』 1.125-135.
- 深井裕二・河合洋明・中野修 (2017) 「学士力分析システムにおけるスキル自己評価ルーブリックの適用」『北海道科学大学研究紀要』 43.1-8.
- 松本剛次 (2021) 「国内の大学における「学士力」の教育としての日本語教育の現状と課題-教学マネジメント上の位置づけと実践事例の検討から-」『日本語教育』 179.109-123.
- 山田嘉徳・森朋子・毛利美穂・岩崎千晶・田中俊哉 (2015) 「学びに活用するルーブリックの評価に関する方法論の検討」『関西大学高等教育研究』 6.21-30.
- 渡辺博芳・新井正之・佐々木茂・盛拓生・古川文人・水谷晃三・眞坂美江子・塩野目剛亮・高井久美子・有本泰子 (2019) 「汎用的能力評価のためのルーブリックとチェックリストの提案」『情報教育シンポジウム論文集』 30-37.

【参考資料】

- 金沢大学 「「人類の未来をつくる独創的研究者の育成」プログラム」 <https://gsc.w3.kanazawa-u.ac.jp/> (2021年8月18日)
- 東北福祉大学 「ルーブリック」 <https://www.tfu.ac.jp/students/rubric.html> (2021年8月18日)