

総合型選抜・学校推薦型選抜を見据えたパフォーマンス評価の研究

——パフォーマンス課題の実践とルーブリックの分析——

中切 正人, 橋本 康弘 (福井大学), 宮下 伊吉 (三重大学), 大久保 貢 (福井大学)

本研究は、パフォーマンス課題とその評価を国立大学の総合型選抜・学校推薦型選抜に活用して、汎用的な選抜・評価方法の確立を目指す実証的研究である。高校生のグループワークによるパフォーマンス課題の実践で測定されるコミュニケーション力の構成要素とその評価規準・評価基準をルーブリックで示した。そして、中切ほか(2019)の実践で信頼性が低く受講者の個人差を捉えるのに適切な規準設定ではなかった構成要素を検討した結果、1つを保留し、他の3つの改善を試みた。その結果、改善版「共同創作力」と言語的観点と測定・評価する改善版「表現力②」の信頼性と受講者の個人差を捉える規準設定が向上した。

キーワード：高大接続, 大学入試, パフォーマンス評価, ルーブリック, ゼミナール教育

1 本研究の目的と背景

本研究は国立大学における総合型選抜・学校推薦型選抜(現行AO入試・推薦入試)を念頭に置いた実践研究である。近代大学の創設者フンボルトの理念「文系ゼミナール教育(理系は実験室教育)」(潮木, 2004)に基づく2017年12月に実施したパフォーマンス評価(中切ほか, 2019)に対し改善を試みる。

1.1 本研究の目的と方法

本研究は文系ゼミナール教育に対応した、グループワークにおけるコミュニケーション力の評価のあり方を分析・検証・考察することを目的とする。それは、中切ほか(2019)で取り上げた2017年12月実施のパフォーマンス課題「未来のカリキュラム」に対するパフォーマンス評価の考察において浮上した、以下の3つの課題を解決することである。

課題(1): グループワークにおけるファシリテーターの活用の意義と、そこで測定される「進行調整力」の測定結果を分析して再検討する。

課題(2): 「共同創作力」のルーブリックにおける評価規準と評価基準を再検討する。中切ほか(2019)で取り上げた2017年12月の「共同創作力」の評価規準は4項目の観点の記述で構成されていた。4観点の記述は一見明確に見えるが、評価者から見ると、それぞれの項目の判定自体にブレや迷いが生じやすい。この4項目の観点の混在が評価の信頼性と受講者の個人差を捉えるのに適切な規準設定に影響を与えた可能性が高いと考え、それに代わる新しい評価規準と評価基準を設定する。

課題(3): 「表現力①」と「表現力②」の違いを明確

にして書き改めることにより、新しい評価規準と評価基準を再検証する。すなわち、プレゼンテーション場面の評価の観点と非言語的観点と言語的観点の二つに絞り、それぞれの評価規準と評価基準を書き改めて、その信頼性と受講者の個人差を捉えるのに適切な規準設定の改善を図る。

2 パフォーマンス評価の改善

最初に中切ほか(2019)に示したパフォーマンス課題と評価の概要を示す。次に、2018年度と2019年度に同じ課題で4回実施した改善版パフォーマンス評価の信頼性と個人差を捉える規準設定を検証する。

2.1 中切ほか(2019)のパフォーマンス課題と評価

2.1.1 中切ほか(2019)のパフォーマンス課題の概要

【表1】は、中切ほか(2019)で取り上げた2017年12月のパフォーマンス課題「2050年の未来の中学校のカリキュラム」プロジェクトの日程とその評価場面である。福井県内の高等学校9校から1年生27名、2年生32名、3年生5名の合計57名(男17名、女40名)の高校生と、7名の教員が参加した。評価者は大学から5名が加わり12名となった。

プロジェクトの朝、受講者から『未来の学校づくりに関する調査研究報告書』(国立教育政策研究所, 2013)の抜粋部(全14頁)を要約した事前課題が提出され、後日これを元に「理解力」が測定・評価された。次の講義では、この事前課題と別の資料(日本交通計画協会自主研究会, 2011)を元に2050年の未来が描かれたスライド、そして、プロジェクト用に創作された未来の学習指導要領やカリキュラム表の作成

表1 パフォーマンス課題の日程と評価場面

開始時刻	内 容	当日評価者が測定・評価するコミュニケーション力	提出物をもとに後日測定・評価された情報収集力と高次の思考力
9:00	【受付】: 提出「事前課題プリント」		理解力
9:30	【開講式】日程と評価者紹介		
9:45	【講義】未来の社会状況と学習指導要領		
10:35	休憩		
10:45	【カリキュラム表の作成(第1段階)】: 提出		適用力 分析力
11:25	【グループ毎のオリエンテーション】自己紹介		
11:50	昼食		
12:50	【グループワーク①】 カリキュラム表の作成(ボードを囲んで)	相互理解力 伝達工夫力 共同創作力 進行調整力	
14:50	【グループワーク②】 プレゼンテーションの準備とリハーサル		
15:30	【プレゼンテーション】個人の発表を評価	表現力① 表現力②	
16:10	【カリキュラム表の作成(第2段階)】: 提出		評価力 創造力
16:50	【閉講式】		

注) 出典は(中切ほか, 2019: 88)。

方法のスライドが紹介された。その後、各自で第1回目のカリキュラム表が作成され、後日「適用力、分析力」が測定・評価された。小休憩の後、グループ毎に各自のカリキュラム表を紹介しあい昼食となった。

午後のグループワーク①では、最初にかリキュラム作成の方針が検討され、カリキュラム表を囲んで話し合いが進んだ。カリキュラムが完成するまでの2時間、評価者により「相互理解力、伝達工夫力、共同創作力、進行調整力」が測定・評価された。次のグループワー

ク②では、プレゼンテーションでの役割分担が決定され模擬練習された。プレゼンでは、一人2分の発表に対して「表現力①、表現力②」が測定・評価された。終了後は、ここまでの活動を総括しつつ各自の2050年のカリキュラム表が作成された。後日これを元に「評価力、創造力」が測定・評価された。

評価者による評価を整理すると、まず、事前課題と当日作成されたカリキュラム表を元に「情報収集力: 理解する、適用する」と「高次の思考力: 分析する、評価する、創造する」が測定・評価された。これらは、アンダーソンらの改訂版タキノミー (Anderson, 2001) の認知過程次元に相当し、また、学習指導要領の「思考力」「判断力」に対応すると考えられる。

次に、グループワークとプレゼンテーションでは「コミュニケーション力: 相互理解力、伝達工夫力、共同創作力、進行調整力、表現力①②」が測定・評価された。これらは、学習指導要領の「表現力」「主体的に学習に取り組む態度」に対応すると考えられる。

2.1.2 中切ほか (2019) での評価とその課題

中切ほか (2019) で測定された2017年12月段階のコミュニケーション力の構成要素の評価基準を下に記す。【表2】はその評価基準で「3」を基準とする。

「相互理解力」: 他者の意見に耳を傾けて合理的な理解・判断を下したり、共感したりする能力。

「伝達工夫力」: 他者の意見に対し論理的な質問をし

表2 2017年12月のコミュニケーション力の構成要素とその評価基準

段階 構成要素	1	2	3(基準)	4
相互理解力	・他者の話す内容を最後まで集中して聞くことができない。あるいは、他者の話の内容に対して合理的な理解・判断を下したり、共感したりすることができない(同意する発言や、何らかの質問をするような反応が見られない)。	・他者の話す内容を最後まで集中して聞くことができ、それに対して合理的な理解・判断を下したり、共感したりすることができる(他者の話の内容に同意する発言が見られたり、もしくは内容に対する何らかの質問が見られたりする)。	・他者の話す内容を最後まで集中して聞くことができ、それに対して合理的な理解・判断を下したり、共感したりすることができる(他者の話の内容のポイントを示すなどして他者の考えに的確に同意する発言が見られたり、話の内容に沿った的確な質問が見られたりする)。	・他者の話す内容を最後まで集中して聞くことができ、それに対して合理的な理解・判断を下したり、共感したりすることができる(他者の話の内容のポイントを示すなどして他者の考えに的確に同意する発言が見られたり、話の内容に沿った的確な質問を通じ、相手の話しやすい状況を作り出し、グループで議論する土台を形成することができる)。
伝達工夫力	・自分の考えや質問事項を、分かりやすく整理することができないため、表情に乏しくあいまいな口調になりがちで、聞き手に分かりやすく伝えることができない。	・自分の考えや質問事項を、分かりやすく整理しているため、表情に乏しくあいまいな口調ながらも、聞き手に分かりやすく伝えることができる。	・自分の考えや質問事項を、筋道立てて分かりやすく整理して、豊かな表情とはっきりとした口調で、聞き手に分かりやすく伝えることができる。	・自分の考えや質問事項を、筋道立てて分かりやすく整理して、豊かな表情とはっきりとした口調で、聞き手に分かりやすく伝えることができる。さらに、相手から質問を受けやすい雰囲気や状況を作り出して、グループで議論する土台を形成することができる。
共同創作力	・4つの観点のうち、いずれか1つを行うことができる。	・4つの観点のうち、いずれか2つを行うことができる。	・4つの観点のうち、いずれか3つを行うことができる。	・4つの観点のうち、すべての観点を行うことができる。
進行調整力	・4つの観点のうち、いずれか1つを行うことができる。	・4つの観点のうち、いずれか2つを行うことができる。	・4つの観点のうち、いずれか3つを行うことができる。	・4つの観点のうち、すべての観点を行うことができる。
表現力①	・発表者の話術がプレゼンテーションの理解を妨げており、発表者が落ち着きなく見える。	・発表者の話術がプレゼンテーションを理解可能にしているが、発表者が自信なさげである。	・発表者の話術がプレゼンテーションを興味深くしており、発表者が落ち着いて見える。	・発表者の話術がプレゼンテーションを積極的にしており、発表者が洗練され、自信があるように見える。
表現力②	・発表者の中心的なメッセージが推測される程度で、聴衆にはメッセージが明確に伝わっていない。	・発表者の中心的なメッセージは基本的に理解可能であるが、その繰り返しが無いことから、記憶に残らない。	・発表者の中心的なメッセージが明確であり、さらに、サポートされる資料(例示、統計、図解、引用、比喩など)によって一貫性を持っている。	・発表者の中心的なメッセージが説得力を持っており、さらに、サポートされる資料(例示、統計、図解、引用、比喩など)によって発表者の信頼性や権威が確立されている(正確に述べられ、適切に繰り返され、記憶に残る)。

注) 出典は(中切ほか, 2019: 86)。

たり自分の考えを伝える工夫をしたりする能力。

「共同創作力」：他者と共同して一つの研究成果をまとめ、完成させる能力で、以下の 4 観点で構成。
 ①メンバーに敬意を払い礼儀正しく振る舞う。②明るい話し方などで前向きな姿勢を伝える。③メンバーの能力を信頼し課題達成の確信を伝えてやる気を高める。④メンバーを手助けしたり励ましたりする。

「進行調整力」：メンバーから意見をうまく引き出し、最終的な合意を導くファシリテーションにかかわる能力で、以下の 4 観点で構成。①意見やアイデアを出しやすいムード作りをする。②意見や議論をまとめて合意形成を促す。③議論すべき内容へ誘導する。④話し合いの合理性を高める。

「表現力①」：プレゼンテーション時に、話術（姿勢・声の表現等）を駆使し、メッセージを聴衆に伝達できる能力。

「表現力②」：プレゼンテーション時に、発表者が中心的なメッセージを聴衆に伝達する能力。

【表 3】には、評価者による評点の占有率（4 段階）と評点差の度数分布と評点の平均値、および「二次の重み付き κ 係数」を示した。評定者間の信頼性を検証する κ 係数は、見かけ上の一致（観察値）から偶然による一致（周辺度数から期待値を使って計算）を差し引いたものを使って計算される。完全一致で 1、偶然の一致と考えられるとき 0、偶然より一致度が低いとき負の値となる（下井・谷, 2007: 130）。一般的に κ 係数 0.21 ~ 0.40 は「まずまずの一致」、0.41 ~ 0.60 は「中等度の一致」、0.61 ~ 0.80 は「かなりの一致」、0.81 ~ 1.0 は「完全一致」とされる。順序尺度では多くの場合「二次の重み付き κ 係数」が使用されるので、これ以降の記述においては、特に断らない限り κ 係数とは二次の重み付き κ 係数を意味する。

【表 3】に示す様に、中切ほか（2019）の「相互理解力」「伝達工夫力」の κ 係数は共に「中等度の一致」以上で、ある程度の信頼性が確認された。また、両者の各評点の占有率は評点「1」の能力が発揮されない評価が見られず、「2」のやや能力が認められる段階が 4 分の 1、基準とした「3」のある程度能力が認められる段階が半分以上で、発揮されている「4」段階が 5 分の 1 認められる。いわゆる両極端に偏らない山型のパターンである。さらに、両者は共に評点差「0」が 50%を超え、評点の平均値が共に 2.9 である。よって、この 2 つの構成要素は受講者の個人差を捉えるのに適切な規準設定であったと考えられる。

これに対し、「共同創作力」の κ 係数は偶然よりも

表 3 2017 年 12 月のプロジェクトの能力別評点の占有率と一致率（%）と平均値および二次の重み付き κ 係数

各指標 能力	12月16日 各評点の占有率(%)				評点差の割合(%)			評点の 平均値	二次の 重み付 き κ 係数
	1	2	3	4	0差	1差	2差		
相互理解力	0	26	54	19	55	45	0	2.9	0.56
伝達工夫力	0	29	51	19	73	27	0	2.9	0.68
共同創作力	0	43	41	16	27	64	9	2.7	-0.09
進行調整力	6	50	26	18	64	18	18	2.6	0.23
表現力①	0	28	52	20	58	33	8	2.9	0.33
表現力②	0	78	19	3	58	42	0	2.2	0.17

注) 出典は(中切ほか, 2019: 90)。

一致度が低いマイナス値の -0.09 であり、「進行調整力」の κ 係数も 0.23 と低いことから、両者の信頼性は低い。また、両者は評点差「2」のケースが存在し、両者の評点の占有率は低い方に偏り、両者の平均値も低いことから、両者共に受講者の個人差を捉えるのに適切な規準設定ではなかったと考えられる。

その理由の考察にあたり、両者の評価規準が 4 観点を記述されている点に注目した。4 観点を記述は一見合理的に見えるが、4 観点はそれぞれに抽象的かつ独立性が高かったことから 1 観点毎の判断の微妙な相違が 4 倍の開きに拡張され、その結果評点の一致度も下がったものと推測される。2017 年 12 月の評価者からも評価しにくい項目と批評された。

次に、プレゼン時の「表現力①」「表現力②」を検証する。「表現力①」は κ 係数が 0.33「まずまずの一致」で、「相互理解力」「伝達工夫力」の両者に次いで信頼性が高く、また、評点の分布や平均値などから判断すると両者に次いで受講者の個人差を捉えるのに適切な規準設定であったと考えられる。これに対し「表現力②」の κ 係数は 0.17 と信頼性が低く、評点の度数分布も下位の「2」に極端に偏り、平均値も 2.2 と極めて低く受講者の個人差を捉えるのに適切な規準設定ではなかったと考えられる。

以上より、2018 年度以降は「相互理解力」と「伝達工夫力」の評価規準と評価基準をそのまま継続して測定し、評価に活用する。しかし、他の 4 つの構成要素は評価規準と評価基準を見直し、その信頼性と受講者の個人差を捉えるのに適切な規準設定に修正する必要があると考え、本研究の課題に設定した。

2.2 2018・2019 年度のパフォーマンス評価の改善

2.2.1 「進行調整力」とファシリテーター

本研究の母体となる 2016 年度のプロジェクトは、

本研究とは異なるパフォーマンス課題であったが、そこに参加した評価者の講評の中に、パフォーマンス課題の遂行に当たりグループの中でファシリテーターを指名したり、別の機会に高校生以外のファシリテーターを養成して活用したりする必要性が指摘されていた(中切ほか, 2017: 179)。しかし、本研究の主執筆者はグループ毎に配されたファシリテーターの個性によってグループワークに差が生まれ、その差が評価の公平性を脅かすのではないかと懸念した。そのため、中切ほか(2019)のプロジェクトにおいても、グループワークの際には一人の進行役(主執筆者)がグループ全体を対象にファシリテートする方法を踏襲した。

しかしながら、ファシリテーターの必要性に迷った本研究の主執筆者は、中切ほか(2019)のプロジェクトでは、グループワーク時にファシリテートする能力を各グループ内のメンバー全員に期待し、その能力の測定を試みた。その際、この能力をグループワーク時の「いつ」「どの場面で」「どの様に」測定するのか、という様な配慮を施すべき事前研究が不十分だった。その結果、評価者は2時間のグループワークの間に「相互理解力」以下の4能力を同時に測定することになり、反省会では測定能力の多さが批判された。その一方で、グループワーク開始時点のファシリテーターの必要性が指摘された(中切ほか, 2019: 89)。

この様に、グループワーク時のファシリテート方法やファシリテーターのあり方の不備が原因で、中切ほか(2019)のプロジェクトでは評価者の「進行調整力」の測定を混乱させることになった。そこで、2018年度はファシリテートのあり方が十分研究されるまでは「進行調整力」の測定と評価を保留することにした。その結果2018年度は、グループワーク時に頻繁に進行役(本研究の主執筆者)と評価者が協力し合っグループの活動状況を情報交換すること、かつ、進行役はすぐにその内容をグループ全体に向かって紹介するという方法で進めることにした。

そして、2019年度は「進行調整力」の測定と評価を保留にするが、グループワークの活性化を目指して外部からファシリテートする方法を試みてきた。それは、同一人物が複数回ファシリテートを担当することによってファシリテート方法の改善と安定化を意図するもので、その担当者を学生・院生に求めた。そして、評価の安定化のためにも同一人物が複数回評価を担当することが必要ではないかと考え、ファシリテーターが同時に評価も担当することを試みてきた。

また、この新たな試みの成果を上げるべく、パフォーマンス課題とそのファシリテートおよびパフォーマン

表4 2018.6.23と9.29と10.13および2019.06.16のプロジェクトの受講者数と2人の評価者による「表現力①, ②」の測定対象の人数

実施日	学校名	受講者全体				2人の評価者による「表現力①, ②」の測定対象			
		学年	男	女	合計	学年	男	女	合計
2018年 6月 23日	A高校	3	0	5	5	3	0	3	3
	A高校	2	1	1	2	2	1	0	1
	B高校	2	2	2	4	2	0	2	2
	C高校	2	5	4	9	2	2	2	4
	D高校	2	2	5	7	2	1	3	4
	合計		10	17	27	合計	4	10	14
9月 29日	E高校	2	3	0	3	2	3	0	3
	F高校	2	4	1	5	2	4	1	5
	G高校	2	0	1	1	2	0	1	1
	合計		7	2	9	合計	7	2	9
10月 13日	H高校	2	3	3	6	2	3	3	6
		合計		3	3	6	合計	3	3
2019年 6月 15日	B高校	2	2	1	3	3	2	1	3
	C高校	2	2	3	5	2	2	3	5
		合計		4	4	8	合計	4	4

ス評価についてまとめた「ファシリテーターと評価者用のマニュアル」を作成して、改訂を重ねてきた。そして、プロジェクト前には必ずそのプロジェクトの担当ファシリテーター全員(学生・院生)を招集し、マニュアルを元にプロジェクト受講者(高校生)に配布した事前課題、パフォーマンス課題やグループワーク・プレゼンテーションのファシリテートのあり方、さらにルーブリック評価等について研修している。

2.2.2 「共同創作力」の改善と検証結果

本研究が分析対象とするパフォーマンス課題「2050年の未来の中学校のカリキュラム」の実践と評価は、2018年6月23日と9月29日と10月13日、および2019年6月15日に実施された。【表4】に、各プロジェクトの実施日別に参加した福井県内の高等学校と受講者数、さらに2人の評価者が測定することができた「表現力」の対象人数を記す。

中切ほか(2019)のプロジェクトの反省を受け、「共同創作力」の評価規準と評価基準を全面的に見直した。まず、中切ほか(2019)には4観点で記述していた評価規準を改め、「相互理解力」や「伝達工夫力」と同じ記述形式にした。以下、変更後の評価規準を記し、その評価基準は【表5】に示す。

「共同創作力」: グループワークにおいて、他者と共

表 5 2018 年度の改善版「共同創作力」および 4 観点の改善版「表現力①」と改善版「表現力②」の評価基準

	1	2	3(基準)	4
共同創作力	・自分の意見やアイデアを一方的に主張し、グループのメンバーを説得して、自分の意見やアイデアに従わせようとする。	・自分の意見やアイデアとグループのメンバーの意見やアイデアを表面的にすり合わせるレベルにとどまる。そのため、相手の理解を十分得られることなく、自己の見解をグループの意見やアイデアとしてまとめようとする。	・自分の意見やアイデアとグループのメンバーの意見やアイデアを互いにすり合わせることで、グループとして一つの意見やアイデアにまとめることができる。	・自分の意見やアイデアとグループのメンバーの意見やアイデアを、深く分析したり比較したりすることによって互いの長所を生かすことに成功し、ワンランクレベルが上の意見やアイデアとしてまとめることができる。たとえば、グループワーク時に議論を深める質問をしたり、相手の見解を正確に理解した上で自己の見解との整合性を図ったり、あるいは修正案を提示したりすることによって、自己と相手の見解を高次元で融合させることができる。
表現力①	・4つの観点のうち、いずれか1つを行うことができる。	・4つの観点のうち、いずれか2つを行うことができる。	・4つの観点のうち、いずれか3つを行うことができる。	・4つの観点のうち、すべての観点を行うことができる。
表現力②	・プレゼンテーション時において、キーワードを適切に使用したり、根拠を明確に示すことができないため、発表内容を論理的に伝えることができない。	・プレゼンテーション時において、キーワードを適切に使用することはできるが、根拠を示した上で論理的に発表内容を伝えることができないため、視聴者の理解を十分得ることができない。	・プレゼンテーション時において、キーワードを適切に使用し、根拠を示したうえで論理的に発表内容をつたえることができるため、視聴者は発言内容を理解することができる。	・プレゼンテーション時において、キーワードを正確かつ適切に使用すばかりでなく、発表内容の根拠を的確かつ明確に示したうえで、理路整然と論理的に説明することができる。そのため、視聴者は発表内容を十分理解し、納得することができる。

同して一つの研究成果（作品）をまとめ、完成させることができる能力。たとえば、グループワーク時に議論を深める質問をしたり、相手の見解を正確に理解した上で、自己の見解との整合性を図る提案や見解を提示したりすることができる能力。

この改善版「共同創作力」の評価規準と評価基準を検証することが出来たのは、全部で 27 人の受講者が参加した 6 月 23 日のプロジェクトであった¹⁾。このプロジェクトではグループワーク時の「共同創作力」を 2 人の評価者で評価することができた受講者は 10 人で、全体の 37% に当たるため分析対象とする。

中切ほか (2019) の 4 観点の記述による「共同創作力」の κ 係数は -0.09 であったが、【表 6】に示す様に 6 月 23 日の改善版では κ 係数が 0.51 「中等度の一致」に向上した。ここで本研究は κ 係数の有意水準を 5% に定める。【表 6】に示す様に、改善版「共同創作力」の両側 p 値は 0.03 であることからこの κ 係数 0.51 は有意である。また、この数値はグループワーク時に測定した「相互理解力」の 0.48 と「伝達工夫力」の 0.59 に匹敵する。よって、改善版「共同創作力」の評価規準と評価基準の信頼性は向上したと考えられる。なお、【表 6】に示す「相互理解力」と「伝達工夫力」の κ 係数の両側 p 値もそれぞれ 0.03 と 0.02 である。よって両者の κ 係数も有意である。

さらに、【表 6】で改善版「共同創作力」の各評点の占有率を見ると、評点「1」の様な能力が発揮されない評価が見られず、評点「2」のやや能力が認められる段階が 2 割で、基準点「3」のある程度能力が認められる段階が 5 割半を占め、「4」の発揮されている段階が 2 割半認められる。いわゆる山型のパターンを示した。また、評点差「2」が無く、評点の平均値も「相互理解力」「伝達工夫力」と共通の 3.1 である。以上の結果は、同時に測定された「相互理解力」と「伝

表 6 2018.6.23 のプロジェクトの能力別評点の占有率と一致率と平均値、および二次の重み付きカッパ係数

6月23日 各指標 能力	各評点の占有率 (%)				評点差の割合 (%)			評点の平均値	二次の重み付き κ 係数	両側 p 値
	1	2	3	4	0差	1差	2差			
相互理解力	0	20	50	30	40	60	0	3.1	0.48	0.03
伝達工夫力	0	25	45	30	50	50	0	3.1	0.59	0.02
共同創作力	0	20	55	25	50	50	0	3.1	0.51	0.03

達工夫力」に近似している。よって、改善版「共同創作力」は、受講者の個人差を捉えるのに適切な規準設定に近付いたと考えられる。

2.2.3 「表現力①」と「表現力②」の改善と検証結果

中切ほか (2019) のプロジェクトの反省を受けて、2018 年度は「表現力①」と「表現力②」の評価規準と評価基準を全面的に見直した。それは、中切ほか (2019) の表現力の評価規準が、共に「聴衆に伝達出来る能力」とされた上で、「表現力①」が「話術」について、「表現力②」が「メッセージ」について記されていたところを、より明確に記述し直したことである。すなわち、改善版「表現力①」は「非言語的表現」にかかわる能力、改善版「表現力②」は「言語的表現」にかかわる能力であると規定した。そして、その上で評価規準と評価基準を書き改めた。以下、変更後の評価規準を示し、その評価基準は【表 5】に示す。

「表現力①」: プレゼンテーション時における非言語的表現にかかわる能力で、以下の 4 つの観点を発揮することができる能力。①自信をもって伝えている。②ジェスチャーが豊かである。③アイコンタクトができていいる。④発声がしっかりして聞

き取りやすい。

「表現力②」：プレゼンテーション時における言語的表現にかかわる能力である。具体的には、キーワードを適切に使用することができ、かつ、根拠を示した上で論理的に発表内容を伝えることができる能力である。

改善版の「表現力①」と「表現力②」の測定に当たり、1人の受講者に対し2人の評価者で評価することができたプロジェクトとその受講者数は以下の通りである（【表4】参照）。すなわち、2018年6月23日の27人の受講者の半数の14人と²⁾、同9月29日の受講者総数の9人と、同10月13日の受講者総数の6人、および2019年6月15日の受講者総数8人の合計37人である。

最初に改善版「表現力②」を検証する。まず、中切ほか（2019）の「表現力②」の κ 係数は0.17だったが、【表7】に示す様に改善版「表現力②」の κ 係数は0.38に向上した。この数値は2018年6月23日に測定した「相互理解力」の0.48、「伝達工夫力」の0.59、改善版「共同創作力」の0.61に準ずる（【表6】参照）。また、【表7】に示す様に改善版「表現力②」の κ 係数の両側p値は0.01であることから有意である。よって、改善版「表現力②」の評価規準と評価基準の信頼性は中切ほか（2019）の「表現力②」よりも向上したと考えられる。

さらに【表7】で改善版「表現力②」の各評点の占有率を見ると、評点「1」の評価が見られず、評点「2」が1割弱で、基準点とする評点「3」が6割半を占め、評点「4」が3割弱認められる。いわゆる山型のパター

表7 2018.6.23と9.29と10.13および2019.06.15のプロジェクトの「表現力②」の評点の占有率と一致率と平均値、および二次の重み付きカッパ係数

各指標 能力	各評点の占有率(%)				評点差の割合(%)			評点の平均値	二次の重み付き κ 係数	両側 p値
	1	2	3	4	0差	1差	2差			
表現力②	0	7	65	28	62	38	0	3.2	0.38	0.01

ンに近付いた。また、評点差が「2」に広がることは無かった。そして、評点の平均値は6月23日のグループワーク時の「相互理解力」「伝達工夫力」と改善版「共同創作力」の3.1に近い3.2であった。以上より、改善版「表現力②」は中切ほか（2019）の「表現力②」よりも、受講者の個人差を捉えるのに適切な規準設定に近付いたと考えられる。

次に、改善版「表現力①」を検証する。改善版「表現力①」の κ 係数の測定は4観点版の「2018年6月23日と9月29日」、そして3観点版の「2018年10月13日と2019年6月15日」の二つのプロジェクトに分けられる。この違いは、前者のプロジェクトで複数の評価者から受けた指摘、すなわち前者の評価規準にある「自信をもって伝えている」観点が他の3観点到比べて抽象的なため、評価基準の判断に迷いやすいという指摘に対処したものである。その結果、後者では【表8】に示す様に、その観点を削った3観点の再改善版「表現力①」に再修正した。

改善の順番に沿って、改善版4観点の「表現力①」から検証する。中切ほか（2019）の「表現力①」の κ 係数は0.33だったが、【表9】に示す様にその κ 係数は0.25となりその信頼性は低下した。ただし、この κ 係数の両側p値は0.21で有意ではない。今後は、3観点の再改善版「表現力①」の測定結果と合わせて、よりよい評価規準と評価基準を作成する際の参考資料として活用したい。続けて、各評点の占有率を見る。評点「1」がわずかに見られ、評点「2」と基準点の評点「3」が共に4割程度を占め、評点「4」が2割弱認められ、この分布は山型のパターンとは異なる。

表9 2018.6.23と9.29のプロジェクトの「表現力①」の評点の占有率と一致率と平均値、および二次の重み付きカッパ係数

各指標 能力	各評点の占有率(%)				評点差の割合(%)			評点の平均値	二次の重み付き κ 係数	両側 p値
	1	2	3	4	0差	1差	2差			
表現力①	2	41	39	18	35	56	9	2.7	0.25	0.21

表8 3観点の改善版「表現力①」の評価規準と評価基準

	(評価規準：測定する内容)	評価基準 (3を基準とする)			
		1	2	3	4
表現力①	・プレゼンテーション時における非言語的表現にかかわる能力で、以下の3つの観点を発揮することができる能力。 ①ジェスチャーが豊かである。 ②アイコンタクトができていいる。 ③発声がしっかりと聞き取りやすい。	・左に示した3つの観点のいずれも行いうことができない。	・左に示した3つの観点のうち、いずれか1つを行うことができる。	・左に示した3つの観点のうち、いずれか2つを行うことができる。	・左に示した3つの観点のうち、すべての観点を行うことができる。

また、評点差が「2 差」のケースが 1 割弱存在し、その評定平均値は 2018 年 6 月 23 日の「相互理解力」「伝達工夫力」と改善版「共同創作力」および改善版「表現力②」の 3.1 よりもかなり低い 2.7 である。よって、4 観点の改善版「表現力①」は受講者の個人差を捉えるのに適切な規準設定ではなかったと考えられる。

続いて、3 観点の再改善版「表現力①」について検証する。まず、信頼性については【表 10】に示す様にその K 係数は 0.36 で、4 観点の改善版「表現力①」の 0.25 よりも向上した。しかし、この K 係数の両側 p 値は 0.15 で有意ではない。今後は、3 観点の再改善版「表現力①」の測定結果と合わせて、よりよい評価規準と評価基準を作成する際の参考資料として活用したい。続けて、各評点の占有率を見る。評点「1」が見られず、評点「2」がわずかに 1 割弱で、基準点の評点「3」が 8 割を占め、評点「4」が 1 割強存在する。この分布は 4 観点の改善版「表現力①」よりも山型のパターンに近い。また、評点差「2」が存在せず、評点差「0」が 8 割を占め、その評定平均値は 2018 年 6 月 23 日の「相互理解力」「伝達工夫力」と改善版「共同創作力」と同じ 3.1 であり、改善版「表現力②」の 3.2 に近い。よって、3 観点の改善版「表現力①」は 4 観点の改善版「表現力①」よりも、受講者の個人差を捉えるのに適切な規準設定に近付いたと考えられる。

表 10 2018.10.13 と 2019.06.15 のプロジェクトの「表現力①」の評点の占有率と一致率と平均値、および二次の重み付きカッパ係数

各指標 能力	各評点の占有率 (%)				評点差の割合 (%)			評点の 平均値	二次の 重み付き K 係数	両側 p 値
	1	2	3	4	0差	1差	2差			
表現力①	0	4	82	14	79	21	0	3.1	0.36	0.15

3 今後の課題

本研究は、中切ほか (2019) で浮上したコミュニケーション力にかかわる 3 つの課題に対処してきた。最初に、「進行調整力」はその測定と評価を保留し、2019 年度からはグループワークの活性化を目指して、専属のファシリテーターが外部からファシリテートし、評価の安定化のためにファシリテーターが評価も担当することを試みている。

次に、2017 年度の「共同創作力」の 4 観点の評価規準と評価基準を改め、グループワーク時に測定される「相互理解力」「伝達工夫力」と同じ記述形式に統一し、新しい評価規準と評価基準を設定した。そ

の信頼性と受講者の個人差を捉えるのに適切な規準設定であったかどうか検証したところ、「相互理解力」「伝達工夫力」に匹敵するレベルに改善された。

最後に、「表現力①」と「表現力②」は、プレゼンテーション場面を評価する視点を非言語的観点と言語的観点の二つに絞り、それぞれの評価規準と評価基準を設定し直した。まず、言語的観点を測定・評価する改善版「表現力②」についてその信頼性と受講者の個人差を捉えるのに適切な規準設定であったかどうか検証したところ、グループワーク時に測定する「相互理解力」「伝達工夫力」と改善版「共同創作力」に準ずるところまで向上した。これに対し、2018 年度の 2 回のプロジェクトで検証した 4 観点の改善版「表現力①」の信頼性は低く、受講者の個人差を捉えるのに適切な規準設定ではなかった。そこで、3 観点の再改善版「表現力①」に改め、再度検証した。しかし、その信頼性は低かった。今後は受講者の個人差を捉える規準設定と合わせて、改善と検証を継続していきたい。

さらに今後は、本研究のパフォーマンス課題と評価の改善を繰り返して、ファシリテートのあり方を確立し、パフォーマンス評価の信頼性を高め、受講者の個人差を捉えるのに適切な規準設定を確立することを目指す。そして、国立大学の総合型選抜・学校推薦型選抜に適用可能な選抜・評価のあり方に寄与したい。

注

- 1) 2018 年度以降のプロジェクトでは、受講者 1 人 1 人の活発な活動を促すために 1 グループを 3 人か 4 人で構成した。そして、どのグループも最低 1 人の評価者が担当し、全ての受講者が送付された受講結果を元に自己の成長を確認し、有効活用出来る様に努めた。しかし、2018 年 6 月 23 日のプロジェクトは、例外的に多くの受講者と評価者を集めることができたことから、グループワーク時に測定する能力の信頼性を測定するために 1 グループ 5 人までの構成とし、2 グループにそれぞれ 2 人の評価者を配置することが出来て、合計 10 人が分析対象とされた。これに対し、プレゼンテーション時は基本的に評価者全員で受講者全員の「表現力①」と「表現力②」を測定・評価することが出来た。
- 2) 受講者の多かった 2018 年 6 月 23 日は受講者のプレゼン時間を 2 分間確保し、かつ評価者が集中力を維持出来る様に 6 グループ (27 人) を二分割して実施した。今回分析に使用した 2 人の評価者による測定対象は 14 人 (3 グループ) である。

参考文献

- Anderson, L. W., et al (Eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing*, Longman.
- 国立教育政策研究所 (2013). 『未来の学校づくりに関する調査研究報告書』平成 24 年度プロジェクト研究調査研究報告書 (研究代表者: 工藤文三).
- 中切正人・雨森聡・大久保貢 (2017). 「文系パフォーマンス評価の実証的研究—パフォーマンス課題における思考力とコミュニケーション能力の測定—」『平成 29 年度全国大学入学者選抜研究連絡協議会研究発表予稿集』, 174-179.
- 中切正人・橋本康弘・宮下伊吉・大久保貢 (2019). 「AO・推薦入試を見据えた文系パフォーマンス評価—パフォーマンス課題「未来の時間割」の実践とコミュニケーション力の評価の分析—」『大学入試研究ジャーナル』 29, 87-92.
- 日本交通計画協会自主研究会 (2011). 『2050 年都市ビジョン研究会中間成果報告』
- 下井俊典・谷浩明 (2007). 「遅発性筋痛測定における 4 種類の疼痛測定法の信頼性」『理学療法科学』 22(1), 125-131.
- 潮木守一 (2004). 『世界の大学危機』中央公論社.