

実践報告

オンデマンド型授業において自己評価力を高めるための 協議ワークを取り入れたピアレビューの効果とプロセス

岩田 貴帆¹・柴田 悠²(¹ 京都大学大学院教育学研究科・² 京都大学大学院人間・環境学研究科)

本稿では、学生の自己評価力を高めるための教授法「協議ワークを取り入れたピアレビュー」についてオンデマンド型授業でのよりよい実施方法の示唆を得ることを目的として、京都大学の社会学の教養科目で当該教授法を行い、3つの分析を実施した。まず、当該教授法の前後で、自己評価力が向上した学生がいた一方で低下した学生もいたため、受講生の自己評価力の平均はほとんど変化しなかった。次に、自己評価力が向上した学生は、低下した学生と比較して、その後レポートをより大きく改善できていたことから、自己評価力の向上はレポートの改善に繋がる傾向が確認された。そして、自己評価力が低下した学生における協議ワークのプロセスの特徴として、自己評価の根拠を投稿する手順から逸脱していたり、評価基準の理解が不十分であったりする課題が明らかになった。改善策として、手順の意義を学生に伝えることや手順を構造化したシステムを開発・利用する等の協議ワークの実施方法について検討した。

キーワード: ピアレビュー、自己評価力、レポート課題、オンデマンド型授業

1. 背景と目的

1.1. オンデマンド型授業における学生等の意見交換

新型コロナウイルスの感染拡大を受け、2020年度、全国の大学でオンライン授業が広く実施されている。オンライン授業とは、大学設置基準第25条と、平成13年文部科学省告示第51号及び平成19年文部科学省告示第114号（以下、メディア授業告示）に基づく授業を指す。

オンライン授業は、大学のキャンパスという場所的拘束から自由であるという点において多様な状況に対応できる可能性を秘めており、こうしたことから、コロナ禍の収束後においても学生を学習目標への到達へと導くひとつの手段としてその活用が期待できる。

オンライン授業には、Zoom等の遠隔通信ツールを用いた同時双方向型（メディア授業告示第1号）と、LMS等を用いて教材配信と指導を行うオンデマンド型（メディア授業告示第2号）がある。本稿ではこのうち後者のオンデマンド型に焦点を当てる。オンデマンド型授業の要件の1つとして「当該授業に関する学生等の意見の交換の機会が確保されている」（メディア授業告示第2号）ことが求められている。本稿で報告する実践では、学生同士の意見交換を重視し、「協議ワークを取り入れたピアレビュー」を導入した点に特徴がある。

1.2. 自己評価力を高める「協議ワークを取り入れたピアレビュー」

本稿で論じる対象とする授業は、知識の習得だけでなく、知識を活用する力の育成を学習目標とするオンデマンド型授業である。佐渡島他（2012; 2015）は、学術的文章を作成する技能を習得することを目的としたオンデマンド型授業で、受講生が作成した全8回のレポートすべてに対して、大学院生が個別にフィードバックを返却している。分析結果からは、履修者の学術的文章作成能力が向上し（佐渡島他, 2012）、この科目を履修した年度の学生は、履修していない年度の学生と比較して、後の専門領域の文章作成課題においても優れた文章を作成している（佐渡島他, 2015）など、この実践方法の有効性が示されている。ただし、このようにフィードバックを行う大学院生10名以上を雇用し、訓練を行った上で、授業を展開するには組織的な取り組みを要する。本稿では、授業科目ごとの取り組みとして実施可能なアプローチを検討する。

知識活用力を育成する授業では、知識を活用するような評価課題に学生が取り組むことを通して、知識活用力を身に付けていくことが必要である。このように評価それ自体が学習経験になることを松下（2016）や小野他（2018）は、「学習としての評価」としている。「学習としての評価」においては、学生自身の自己評価が重要とされている（Earl, 2013）。なぜなら、学生が自己評価することで、自分のパ

パフォーマンスの優れた点・改善を要する点を自分で把握できるようになれば、パフォーマンスを適切に改善することが期待できるからである。特にオンデマンド型授業では、対面授業のように他の受講生の学ぶ姿を参考にすることが難しい環境の中で、課題への取り組みを通して学習を進めることが求められるため、学生が自分でパフォーマンスを評価し、その質を高めることは重要といえるであろう。そこで本稿では、「評価基準を理解し、それを自分のパフォーマンスに適用することで、目標とのギャップを適切に把握する力」として定義される自己評価力（岩田・田口, 2020）に着目し、オンデマンド型授業において自己評価力の向上と、学生のパフォーマンス改善について検討していく。

自己評価力を向上させるための教授法として「協議ワークを取り入れたピアレビュー」（岩田・田口, 2020）を挙げることができる。この教授法は、ルーブリックを用いて①プレ自己評価②学生同士（ピア）の評価③協議ワーク④ポスト自己評価という4ステップで構成される。③の協議ワークとは、①プレ自己評価と②ピア評価を独立して実施した上で、それらの得点や根拠を両者が持ち寄って比較し、得点や根拠がズレた理由を話し合うものである。ピアと評価し合うだけの一般的なレビューでは、ピアからの評価を本人が無批判に受容してしまう問題が存在するが（Yucel et al., 2014）、協議ワークを取り入れることによって、自己評価とピアからの評価のどちらかが（あるいは両方が）誤っている可能性を前提としてピアからの評価で本人が重要とみなした部分だけを取り入れることができる。

岩田（2020a）では、教室対面型の授業において、協議ワークの後の④ポスト自己評価は協議ワークの前の①プレ自己評価と比較して教員評価に近づき、その差は統計的に有意であったこと、すなわち自己評価力が高まることを検証した。さらに岩田（2020b）では、自己評価力が高まった学生は、パフォーマンスを適切に改善できることを明らかにした。すなわち、協議ワーク後に適切に自己評価できていない学生と比較して、適切に自己評価できている学生は、その後パフォーマンスをより大きく改善できていた。このことから、「学習としての評価」論を踏まえて期待できる通り、学生の自己評価力が高まると、パフォーマンスを適切に改善できるという経路が示唆されている。

ただし、岩田（2020a, 2020b）の結果は同一教員が担当する類似した2つの授業科目での実践に基づくものであり、過度な一般化はできない。同様の結果を期待するには、少なくとも、事前に評価基準を理解する学習活動を行うことと、学習到達目標・パフォーマンス課題・評価基準が十分に対応していることが前提として必要と考察されている（岩田, 2020a）以外には、どのような文脈的条件の授業で効

果を持つかは実証的に明らかにされていない。特に、本稿で焦点を当てるオンデマンド型授業の文脈において「協議ワークを取り入れたピアレビュー」を実施するには次に述べるような困難を伴うため、これを回避する工夫を実施した上で、効果を検証する必要がある。

1.3. オンデマンド型授業で「協議ワークを取り入れたピアレビュー」を実施する際の困難

教室で実施する場合には、①でワークシートに記入したプレ自己評価は、②のタイミングではレポートのみをピアと交換し、ワークシートを本人の手元に残せば、ピアはプレ自己評価を知ることなく相手への評価を行うことができる。もし、LMSのフォーラム上で学生がすべての活動を投稿すれば、②のタイミングでピアは①の本人のプレ自己評価が見えてしまい、これを参考に（場合によっては相手の自尊心を傷つけないように過度な配慮をして）ピア評価を書き込むと予想される。このような事態になると、協議ワークでプレ自己評価とピア評価を比較する意味が損なわれてしまい、ピアからの評価を通して本人が自己評価を見直すことができず、自己評価力が高まらない恐れがある。

このような困難を回避する工夫を施して、オンデマンド型授業で「協議ワークを取り入れたピアレビュー」を実施すると、自己評価力が高まるのかどうか、自己評価力が高まった学生ほどその後のパフォーマンスを適切に改善できるのかどうかは明らかでない。さらに、よりよい実施方法を検討するための手がかりを得る必要がある。

1.4. 目的

そこで本稿では、オンデマンド型授業において「協議ワークを取り入れたピアレビュー」を実施すると、(I) 受講生全体の傾向として自己評価力が高まるか、(II) 自己評価力の向上はパフォーマンスの改善に繋がるか、(III) 自己評価力が向上した学生・向上しなかった学生それぞれの協議ワークのやりとりの特徴はどのようなものか、を明らかにする。そのことにより、オンデマンド型授業における「協議ワークを取り入れたピアレビュー」のよりよい実施方法の示唆を得ることを目的とする。

2. 対象とする授業の概要

2.1. 授業の概要と到達目標

京都大学全学共通科目（学部横断型の教養科目）「社会学I」の授業実践（2020年度前期開講）を実践のフィールドとする。第二筆者は2016年度より本科目を担当しており、第一筆者は2018年度よりTAを務めている。この両者で協働して、課題レポートや評価ルーブリックの設

計、授業の運営、レポートの採点などを行っている。2019年度までは教室対面型であったが、2020年度はコロナ禍を受けて、オンデマンド型での授業実施となった。

本授業科目は、同内容で授業曜日が異なる2つのクラスが存在する。両クラスのシラバスや授業内容は完全に同じである。受講登録者数は木曜クラス50名、金曜クラス47名であった。学年は1年次生が70%以上であった。

本科目では、学部横断の教養科目であることも踏まえ、社会学の知識を習得すること以上に、社会学に特徴的な考え方を活用する力の育成を重視している。シラバス記載の学習到達目標は、「社会学的思考法を用いて、現代のさまざまな社会現象や自分自身の人生・生活の背景にある『しくみ』（社会構造とコミュニケーションの相互作用）を、基礎的な水準で分析し説明できるようになる」ことである。社会学的思考法とは、「コミュニケーション」と「構造」の相互影響関係に着目する思考方法のことである。ここでいう「コミュニケーション」、「構造」は社会学の用語であり、これらの相互影響関係に着目しながら社会現象の分析・説明を行うことによって、その社会現象の暗黙の前提を指摘したり、それを踏まえた有効な問題解決策を提示したりすることができる。

2.2. 授業科目全体の設計

学習目標への到達を促す授業を設計するために、4つの下位目標を設定した。

- ①社会学とは何かを理解する
- ②社会学的思考法とは何かを理解する

③社会について扱った資料から問いを設定できる

④設定した問いについて社会学的に説明できる

授業科目全体の流れとして、おおまかにこの下位目標①～④の順で学習が進行していくよう配置した(表1)。

主要な学習活動としては、(1)教材を踏まえた学習レポート作成(2)学習レポートに対して学生同士でコメント活動(3)期末レポートのサンプルを用いた評価練習(4)期末レポートの草稿を学生同士で検討する協議ワークを取り入れたピアレビューの4点で構成される。期末レポートに関する学習活動(3)と(4)は、「学習としての評価」(Earl, 2013)を踏まえて、期末レポートに取り組むことが評価のためだけでなくそれ自体が学習となるよう、学生の自己評価力を意図して配置した。岩田(2020a)で考察されているように、まずサンプルレポートを用いた評価練習により評価基準を理解してから、「協議ワークを取り入れたピアレビュー」を実施することで、学生が自己評価力を向上させることが期待できる。そして、自己評価力が向上すれば、レポートの改善すべき点を自分で把握することができるようになり、レポートの質を高めることをねらいとした。以上のねらいから本授業では自己評価力を重視していることは、学習活動(3)(4)の手順の指示の中で学生に説明した。成績評価の40点分は、この4つの学習活動の参加状況によって採点した。残りは、学期末レポートの草稿10点、学期末レポート50点と配分した。

なお、表1中の「授業開始前」とは、正式な授業開始の前の段階(2020年4月中)において、本授業科目の進め方を教員側・学生の双方が確認し、疑問点を解決

表1 授業科目全体の設計

	教材	学習活動	下位目標との対応
授業開始前	・教員の自己紹介を兼ねたテキスト教材 ・柴田悠『子育て支援と経済成長』「はじめに」「第1章」～「第3章」	(1)教材を熟読し、疑問について学習レポート作成 Zoomでの質疑応答に参加(任意)。	なし
第1週	・教員の独自テキスト教材「社会学の定義・意義・背景」	(1)教材を熟読し、疑問について学習レポート作成。 採点結果と個別コメントを確認。 (2)他の学生の学習レポートに対するコメント活動 Zoomでの質疑応答に参加(任意)。	①②
第2～4週	・柴田悠『子育て支援と経済成長』 「第4章」～「第6章」「おわりに」	(1)教材を熟読し、社会学的思考法を用いて学習レポート作成。採点結果と個別コメントを確認。 (2)他の学生の学習レポートに対するコメント活動 Zoomでの質疑応答に参加(任意)。	①②③④
第5週	・ループリックとサンプルレポート4種	(3)サンプルレポートを用いた評価練習	②
第6週	・テキスト教材(『日本経済新聞』に掲載の教員自筆記事)	第1～4週と同じ 期末レポートの草稿の作成	③④
第7週	なし	(4)協議ワークを取り入れたピアレビュー	③④
第8週	・テキスト教材(内閣府会議に教員が提出した資料)	第1～4週と同じ	①②③④
第9週	なし	期末レポートの草稿を改善して、完成稿を提出	③④

し、授業方針の合意形成を行うための期間であった。よって、課題は設定したが、成績評価の対象外とした。

教材の配信や課題の指示はすべて京都大学のLMS「PandA」で行い、重要な連絡は随時学生メールにも送信した。また、補足的にZoomミーティングを毎週実施した。これは教員とリアルタイムで質疑応答のやりとりに参加したい学生のためのものであり、授業参加点には関係しないことを学生に伝えた。学習内容、学習方法に関する説明はLMS上だけでも伝わるよう配慮した。以下、順に（1）～（4）の学習活動の設計について詳述する。

（1）教材を踏まえた学習レポート作成

下位目標①②の理解については、前年度同様、第二筆者の作成した独自教材や文献を教材として用いてオンデマンド配信を行い、知識を伝達することにした。教材のうち『子育て支援と経済成長』は書籍であり配信は困難なため、学生が入手するよう指示した。教室対面型のように学生の理解を確かめながら説明することができないため、テキスト教材を読んだ上で、その理解に基づいてレポートを作成することを学習課題として課した。

理解できなかった部分や浮かんできた問いについて、400字以上のレポートに調べてまとめること（表1の授業開始前・第1週）や、教材の内容から考えられる「構造とコミュニケーションの（一方向のあるいは双方向の）影響関係」について他の文献を1つ以上参照しながら説明すること（第2週以降）を指示した。

この学習課題により、学生らが自分の下位目標①②の理解を確かめたり疑問を解決したりすることを意図した。浮かんできた疑問が解消しない場合には、その旨をレポート内に記載することで、教員よりZoomで解説を行った。通信環境の関係でリアルタイムにZoomに出席できない学生のために、Zoomを録画しておき、LMSにアップロードし、環境の整った状態で見ることを促した。

（2）他の学生の学習レポートに対するコメント活動

学生同士の意見交換の機会として他の学生がLMSのフォーラムに提出した学習レポートに対して2人以上にコメントすることを課題として指示した。その際、学習到達目標と関連付けて課題指示文を作成した。

すなわち、本授業科目の到達目標である、社会学的思考法を活用して社会現象を説明するにあたっては、経験的根拠に基づいて客観的な議論を行うことが重要である。そこで、コメント活動においても、「単なる会話ではなく、できるだけ『社会科学的なディスカッション』として相応しい水準のやりとりを目指してください」と、学習到達目標に関連付けた指示を行った。その例として「同意します。なぜなら～～～」「私は異なる意見です。なぜなら～～～」「関

連して次のようなことを考えました。～～～。」「わからない部分があったので質問します。～～～?」といったパターンの例を併せて伝えた。

「社会学は社会科学の一領域であること」「社会科学とは社会現象の法則性を、経験的根拠に基づいて説明する学問であること」は、第1週の教材にて解説した。

（3）期末レポートのサンプルを用いた評価練習

「協議ワークを取り入れたピアレビュー」を実施する前提として、学習到達目標・パフォーマンス課題・評価基準が十分に対応していることが必要である（岩田, 2020a）。そこでまず、学生の社会学的思考法を活用する力が表出するようなパフォーマンス課題として次のようなレポート課題を設計した。

新聞記事を1つ選んで要約した上で、そこで扱われている社会現象をコミュニケーションと構造の相互の影響関係に着目して分析・説明しなさい。選んだ新聞記事よりも深い説明となるように、内容に関連する文献を他に2つ以上参照すること。

次に、このレポートの質を判断するための評価基準を示すものとして、表2のようなルーブリックを設計した。レベル3の水準が、本授業科目の学習目標に到達した水準となるように意図した。

レポート課題とルーブリックの開発にあたっては、4種の

表2 期末レポート課題の評価用ルーブリック

レベル4	レベル3の水準を満たした上で、一般的な新聞記事やニュースのレベルを越えるような深い認識枠組みを読者に提供するレポートである。
レベル3	社会現象における人々のコミュニケーションを規定する構造を客観的に説明しており、さらに、その構造がどのようなコミュニケーションから形成されるか（コミュ→構造→コミュ）、あるいは、そのコミュニケーションがどのような構造を形成するか（構造→コミュ→構造）を客観的に説明することができている。客観性のために、事例・データや他の文献を効果的に用いている。
レベル2	社会現象における人々のコミュニケーションを規定する構造を客観的に説明できている。しかし、その構造がどのようなコミュニケーションから形成されるか（コミュ→構造→コミュ）、あるいは、そのコミュニケーションがどのような構造を形成するか（構造→コミュ→構造）の説明は不適切である。説明を試みていたとしても、事例・データや他の文献を効果的に用いることができておらず、客観性が低く不適切な説明となっている。
レベル1	社会現象における人々のコミュニケーションを規定する構造を客観的に説明できていない。説明を試みていたとしても、事例・データや他の文献を効果的に用いることができておらず、客観性が低く不適切な説明となっている。
レベル0	そもそも社会現象を説明しようとしたレポートになっていない。

サンプルレポートをTAである第一筆者が作成し、授業担当者である第二筆者が採点し、協議を重ねることによって、ルーブリックの記述語・課題の指示文・サンプルレポートの3要素を調整して完成させた。

下位目標③④については、学生が「やってみる」だけでは不十分であり、「どのような説明ができていれば、社会的思考法を活用した説明として優れているのか」を学生が理解すること、すなわち評価基準の理解が必要不可欠といえる。そこで、第5週の学習課題としてサンプルレポートを用いた評価練習を実施した。これは、学生が評価基準を理解するために、教員が用意したサンプルを学生が評価基準を用いて評価した上で、教員による評価やその根拠の解説を受けるという学習活動（岩田, 2020a; Rust et al., 2003）を指す。前年度はこれをグループ学習で行っていたが、これをLMSでの実施は困難であるため本実践ではルーブリックとサンプルレポートを配信し、Googleフォーム上のワークシートを用いて個人学習として実施した。サンプルレポートは、筆者らがルーブリック開発に際して作成したものをを用いた。ルーブリックのレベル1から4にそれぞれ該当するサンプルを用いて評価練習することで、学生がレベル間の差異を理解することを期待して設計した。

(4) 期末レポートの草稿について「協議ワークを取り入れたピアレビュー」

1.3. で指摘したようなオンデマンド型授業で「協議ワークを取り入れたピアレビュー」を実施する際の困難を回避するため、プレ自己評価とピア評価を独立して実施する手順を考案した（図1）。

まず、期末レポートの草稿を、LMSのフォーラムに書き込むように指示した。併せてルーブリックを用いたプレ自己評価をする必要があるが、プレ自己評価をフォーラムに書き込んでしまうと、ピア評価に影響を与えてしまう。そこで、LMSではなくGoogleフォームに投稿するよう指示した。

次に、Googleフォームへの投稿が完了してからピア評価をLMSのフォーラム上に書き込むよう指示した。本人のレポートを読んだ上で、ルーブリックを見ながら採点して、評価得点とその根拠を書き込んでもらった。

ピアからの評価を確認した本人は、既にGoogleフォームに投稿済みのプレ自己評価を再度、今度はフォーラム上に記入するよう指示した。

以上のように自己評価をいったん別の場所で実施し、教員側に送信しておくことによって、独立性を確保した。協議ワークの進め方について、以下のように指示した。

自己評価とピア評価を比較して、得点や根拠にズレがあるか確認してください。正確に自己評価やピア評価することは難しい行為ですので、ズレが生じることがよくあります。

もしズレがあった場合には、「なぜズレがあったのか」について考えられる理由（評価基準表の〇〇という語をよくわかっていなかった、自分の/あなたのレポートの〇〇の箇所を誤解して読んでいた等）を書き込み、意見の一致を目指して2人で返信を何度かしあってください。やりとりをしても意見が一致しない場合は、無理に意見を合わせる必要はありません。

期末レポートの草稿に対して「協議ワークを取り入れたピアレビュー」を実施することで、より適切に自己評価できるようになり、完成稿の提出に向けてどのように改善すればよいのかを明確にすることが期待される。

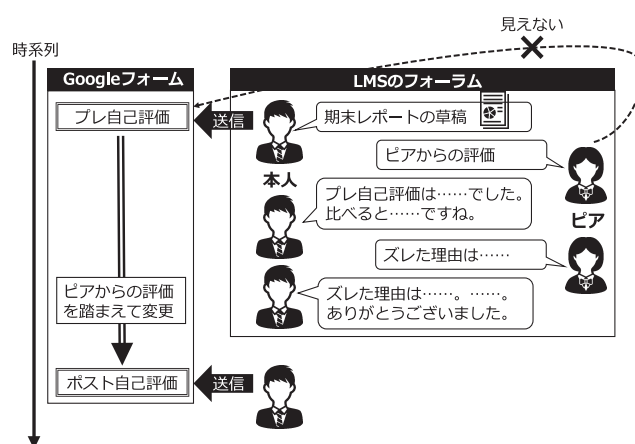


図1 自己評価・ピア評価の独立性を担保する手順

3. 分析方法

研究協力の同意が得られた学生のうち、データ欠損のない68名（木曜クラス33名、金曜クラス35名）を分析対象とする。両クラスで授業の進行に大きな差は生じなかったこと、各データの分布や平均値差を統計的に確認したところ、有意な差は存在しなかったことから、両クラスの受講者を区別せずに分析を行う。量的な分析には統計ソフトR 4.0.0を用いる。有意水準は5%とする。

1.4. 目的で述べたように、本稿の分析を通して明らかにすることの1点目は、オンデマンド型授業において「協議ワークを取り入れたピアレビュー」を実施すると、受講生全体の傾向として自己評価力が高まるか、であった。これに対応して分析Iでは、協議ワークの前後において、草稿レポートに対する自己評価が教員評価に近づいたかどうかを量的に分析する（図2）。自己評価・教員評価の得点はルーブリック得点を用いる。「自己評価と教員評価の差の絶対値」を「自己評価の不適切さ」（岩田・田口, 2020）とする指標を用いる。この指標が0に近い値であるほど、学生は適切に自己評価できていることを意味する。この指標には統計学的厳密性に課題もあるが、当該教授

法の前後で学生が適切に自己評価できるように推移しているかどうかを量的に検討することができる利点がある（岩田, 2020a）。「プレ自己評価の不適切さ」と「ポスト自己評価の不適切さ」を対応あり t 検定で比較し、効果量として標準化平均値差 d （Hedges, 1981）を算出する。両データの間に学生が経験した「協議ワークを取り入れたピアレビュー」が影響して指標が変動することになる。

本稿の分析を通して明らかにすることの2点目は、自己評価力が高まった学生ほどパフォーマンスが改善されるか、であった。これに対応して分析Ⅱでは、まず自己評価力の変動によって学生を3群に分ける。具体的には、協議ワークの前後で、指標「自己評価の不適切さ」が低下した群、変化しなかった学生群、「自己評価の不適切さ」が上昇した群の3群に分ける。この3群それぞれが、草稿に対する「協議ワークを取り入れたピアレビュー」の後に、完成稿提出までにどの程度レポートを改善できたのか、教員得点の変化を群ごとに算出する（図2の分析Ⅱ）。「自己評価の不適切さ」が低下した群、すなわち自己評価力が向上した群の改善幅が大きく、逆に「自己評価の不適切さ」が上昇した群、すなわち自己評価力が低下した群の改善幅が小さければ、協議ワークで高まった自己評価力はその後のパフォーマンス改善に繋がることが確認されたことになる。

本稿の分析を通して明らかにすることの3点目は、自己評価力が向上した学生・向上しなかった学生それぞれの、協議ワークのやりとりの特徴はどのようなものか、であった。これに対応して、協議ワークの各学生のやりとりの投稿記録を、プロセスの特徴によって分類する（図2の分析Ⅲ）。その際、協議ワークにおいて重要と思われるプレ自己

評価の得点とピアからの評価の得点が不一致のケースに着目する。不一致のパターンとしてはA「プレ自己評価は不適切・ピアからの評価は不適切」とB「プレ自己評価は適切・ピアからの評価は不適切」が存在するため、それぞれについて協議ワークのプロセスを分類する。

4. 結果と考察

4.1. 授業科目全体の結果

授業科目は全体的に予定通り進行することができ、学生からは概ね指示通り、各学習活動に取り組んだと見受けられた。

学習レポートは、教材を読んで一定の理解に基づいて作成されているか、疑問点や浮かんできた問いについて調べてまとめているかどうかで成績評価をつけたところ、95%以上がそれに当てはまるものであった。例えば、次の学生236の例では、西欧諸国で1990年代に女性の職場進出が進んだ理由について調べてまとめており、優れた学習レポートの典型例といえる。一方で、取り組み方が不十分な学習レポートについては、TAから個別コメントとして具体的な改善方法を送信した。

学習レポートに対するコメント活動ではコメント活動で学生に指示した「『社会科学的なディスカッション』として相応しい水準のやりとり」を満たすコメントが、授業設計時の予想以上に多く投稿された。ここでいう「社会科学的なディスカッション」とは、第1週の教材において学生らに示した、経験的根拠に基づいて説明することを指している。主観的な印象でコメントを書き込むのではなく、文献を引用するなど、事実やデータに基づいて自らの主張を構成することがそれにあたる。

次の例において、コメントをした学生237は、下線の箇所において、別の文献を用いて議論を発展させている点で、社会科学的なディスカッションといえる。このような水準のコメントが多く見られた。2.2. (2) で述べたように、社会科学においては経験的根拠に基づいて説明することが重要であるという第1週の教材で扱われた学習内容と関連付けて課題指示文を記載したことが、これをもたらしたと考えられる。

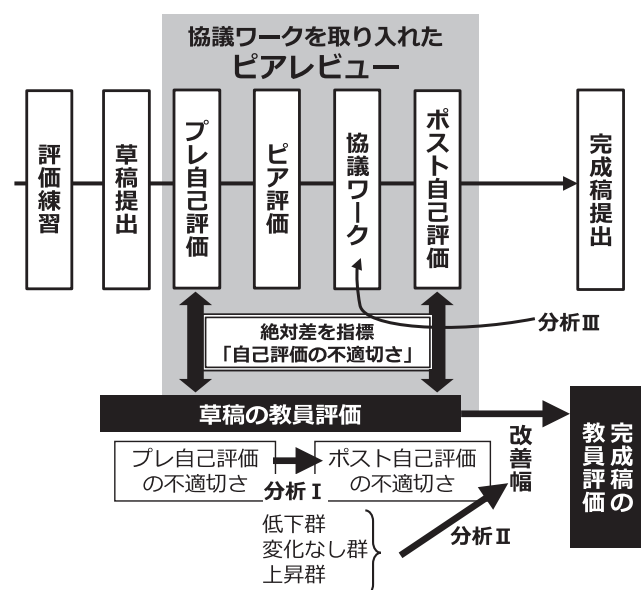


図2 分析に使用するデータと実践の関係

学生236がフォーラムに投稿した学習レポート

課題の新書の140頁で述べられている「ドイツ、フランスなどの西欧諸国では1990年代に女性の職場進出が急速に進んだ」という現象について社会学的に説明する。参考文献2によれば、(略)。したがって先述の現象は「女性が高学歴化したという構造や、女性が育児や家事をしつつも多様な形態で働けるようになったという構造が、女性が積極的に企業に求職するというコミュニケーションにつながった」と言えそうだ。

学生 236 のレポートに対する学生 237 からのコメント

2つ目の構造とコミュニケーションに関して、学生 236 さんの意見に同意します。しかし、1990 年代にドイツで進んだ女性の職場進出は「高学歴の女性が子供を産むことを諦めて進んだ現象」であり、実際に両立支援が注目され始めたのは出生率がなかなか上昇しないことに着目され始めた 2000 年代のようです (*1)。日本と少し状況が似ているように感じたので、日本はドイツを参考に政策を打ち出せるのではないかと思います。(*1) 矢澤朋子 [2013] 大和総研欧州経済「日本の女性就業率は欧州より低い? 注目すべきは就業率の男女差」9-10 頁

上記の学生 236 と 237 のコメント活動は一例として示したものであるが、このような建設的なディスカッションが数多く見られた。このようにコメント活動を通して建設的なディスカッションを行うことに受講生らが慣れていったことが、第 7 週で実施した「協議ワークを取り入れたピアレビュー」でもお互いの意見を尊重したスムーズなやりとりに繋がっていたとみられる。

期末レポートの教員評価（ルーブリック得点）の平均値は 2.53 であった。得点の分布は、1 点が 11 名、2 点が 17 名、3 点が 33 名、4 点が 7 名であった。授業の学習目標への到達を表す水準として設計された 3 点、およびそれを越える 4 点の学生の割合は約 58.8% であった。

分析に関わる各変数の平均・標準偏差を表 3 の上 5 段に示した。また、各受講生のプレ自己評価に対する「自己評価の不適切さ」指標を 3. 分析方法で示した方法で算出し、プレ自己評価の不適切さとして表 3 の下から 2 段目に示した。同様に、各受講生のポスト自己評価に対する「自己評価の不適切さ」指標を算出し、ポスト自己評価の不適切さとして表 3 の最下段に示した。

4.2. 分析Ⅰ 自己評価力の向上についての分析結果

対応あり t 検定により、プレ自己評価の不適切さの平均値 (0.76) とポスト自己評価の不適切さ (0.71) を比較した。その結果、 $t(67) = -0.893, p = .375$ となり有意差はなかった。すなわち、「協議ワークを取り入れたピアレビュー」の前後で自己評価力の平均は有意な向上をしなかった。効果量は $d = 0.11$ であった。

効果量 d の目安は、0.8 で大、0.5 で中、0.2 で小とされている (Cohen, 1988) ことや、教室対面型で実施した同科目では $d = 0.18$ (ルーブリック 4 観点合計得点) であった (岩田, 2020a) ことを踏まえると、受講生の平均的傾向としては「協議ワークを取り入れたピアレビュー」の効果がかなり小さかったといえる。実施方法に何らかの課題がある

表 3 各変数・各指標の基礎統計量

n = 68	平均	標準偏差
プレ自己評価	2.81	0.78
ピアからの評価	2.74	0.82
ポスト自己評価	2.75	0.76
草稿教員評価	2.25	0.95
完成稿教員評価	2.53	0.89
プレ自己評価の不適切さ	0.76	0.79
ポスト自己評価の不適切さ	0.71	0.75

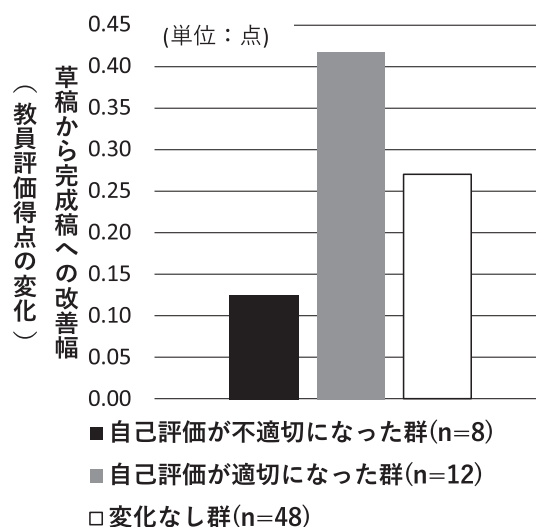


図 3 自己評価の変化とその後のレポート改善の関係

可能性が考えられる。具体的にどのような課題があったのかは、分析Ⅲを踏まえて考察を行う。

4.3. 分析Ⅱ 自己評価力の向上がその後パフォーマンス改善に与える影響の分析結果

「協議ワークを取り入れたピアレビュー」の前後で、指標「自己評価の不適切さ」が下がった学生、すなわち適切に自己評価できるようになった学生は 12 名であった。同指標が変化しなかった学生は 48 名であった。同指標が上がった学生、すなわち、自己評価が不適切になった学生は 8 名であった。この 3 群について、協議ワークを取り入れたピアレビューの後に、最終提出までにどの程度レポートを改善できたのか、草稿から完成稿にかけての教員評価の改善幅を群ごとに算出した (図 3)。サンプルサイズが小さいため平均値差を統計的に検定することはできないが、自己評価が適切になった学生はその後レポートを平均で +0.42 点改善できているのに対して、自己評価が不適切になってしまった学生はその後レポートを平均で +0.12 点しか改善できていない。このことから、オンデマンド型授業

で「協議ワークを取り入れたピアレビュー」を実施した場合でも、教室対面型（岩田, 2020b）と同様、自己評価力はその後のレポート改善に繋がる傾向が見て取れる。

この傾向を踏まえると、もしも「協議ワークを取り入れたピアレビュー」を通して自己評価力が向上した学生がより多ければ、学習到達目標を表す水準以上の期末レポートを書いた学生の割合は、58.8%よりもさらに大きくなっていったと考えられる。

そこで、協議ワークを通して適切に自己評価できるようになった学生と、むしろ不適切になってしまった学生では協議ワークのやりとりによどのような特徴があるのか、プロセスを調べることで、よりよい実施方法のための示唆を得ることとする。

4.4. 分析Ⅲ 「協議ワークを取り入れたピアレビュー」のプロセス

協議ワークは、プレ自己評価とピアからの評価の得点だけでなく根拠についてもズレに着目するものであるが、ここでは特に重要と考えられるプレ自己評価とピアからの評価の得点が不一致であったケース 21 件に着目する。その内訳は、パターン A（プレ自己評価は不適切・ピアからの評価は適切）が 9 件、パターン B（プレ自己評価は適切・ピアからの評価は不適切）が 12 件、であった。これらの LMS 上のやりとりや、Google フォームに提出されたプレ自己評価・ポスト自己評価のデータから、その特徴は、パターン A は 3 つに分類でき、パターン B は 5 つに分類できた（表 4）。以下、それぞれについて詳述する。

(1) パターン A のプロセス

A-1 は、プレ自己評価とピア評価の根拠で不一致だった箇所について焦点化したやりとりの結果、本人は自己評価を適切に修正することができており、「協議ワークを取り入れたピアレビュー」が意図通りに機能したとみられるプロセスである。ここでいう意図とは、本人がピアからの評価を無批判に受容することなく、よく検討した上で必要な部分を自己評価に取り入れるということである。学生 142 は、図 4 の下線①において、すぐにピアからの評価を受け入れるのではなく、自分のプレ自己評価の根拠を示しつつ、ピアが感じたレポートの客観性について再度ピアに訊ねている。その結果、ピアがその箇所については深く検討していなかったことが判明し、より建設的な方向へと議論が進んでいる。

このように、プレ自己評価の得点と根拠を書き込むことによって、プレ自己評価とピアからの評価のいずれか（もしくは両方）が誤っている可能性があるという前提のもと、議論を進めることができるという協議ワークが機能しているのが、A-1 である。A-1 に分類された 4 名は、プレよりも適切にポスト自己評価できている。

ところが、A-2 と A-3 においては、この手順が守られていない。A-2 「プレ自己評価の根拠を書いていないが、ピアからの評価を見て、プレ自己評価の誤りにすぐに気付いた」と A-3 「プレ自己評価を開示しておらず、不一致に着目したやりとりはしていない。とにかくピアの意見を取り入れた」は、プレ自己評価の根拠を（A-3 は得点すら）投稿していない。この 5 名については、ピアからの評価が適切であったから結果としてポスト自己評価は適切になったものの、もしピアからの評価が間違っていたら、不適切なポスト自己評価となっていた可能性がある。そうすると、改善すべき点を見誤ってしまい、完成稿に向けてどのように修正すべきかの判断を誤る恐れがある。

なお、ポスト自己評価においてプレ自己評価から変更しなかったケースはパターン A には存在しなかった。

(2) パターン B のプロセス

プレ自己評価が適切・ピアからの評価が不適切であるパターン B の中で、B-1 は、プレ自己評価の根拠をピアに伝えたところ、ピアが評価の誤りに気付いて改めたので、本人はプレ自己評価の得点と根拠を変更しなかったという、無批判な受容を防ぐことができたプロセスである。その結果、B-1 の学生は適切なプレ自己評価をポスト自己評価でも維持することができた。このような協議ワークが機能したケースは 1 名のみに留まった。

B-2 と B-3 は共に、協議ワークの手順から逸脱したプロセスである。B-2 は、本人がプレ自己評価の根拠を投稿しなかったため、ピア評価が誤っている可能性を両方で検討する機会のないまま、無批判にピア評価を受容してしまっている。協議ワークの課題指示文では、「Google フォームから提出したプレ自己評価の得点と根拠をさらに投稿してください」としているにも関わらず、B-2 の 3 名が投稿したのは得点のみであり、根拠を投稿していなかった。図 4 の学生 224 の Google フォームで提出したプレ自己評価の根拠（図 4 の下線②）は、教員側から見ても妥当である。ピアは、学生 224 は既に新聞記事を引用することで議論の客観性を高めていることを見逃している。にも関わらず、学生 224 はピアが正しいと受容してしまい、ポスト自己評価を 2 に下方修正してしまった。

このようにピアからの評価を無批判に受容してしまう問題は、協議ワークを伴わないピアレビューに存在する問題である（e.g. Yucel et al., 2014）。もし本人が自己評価の根拠を投稿していたならば、ピアからの評価が誤っている可能性も考慮しながら協議ワークを進めることができ、自己評価を適切なまま維持できたかもしれない。

それではなぜ本人はプレ自己評価の根拠を投稿するという手順から逸脱してしまったのだろうか。学生 224 の図

表 4 協議ワークのプロセスを分類した結果と協議ワークの効果・理由の考察

	分類 記号	協議ワークのプロセスの分類	人数	協議 ワーク	プロセスが生じた 理由
プレ自己 評価= 不適切・ ピアから の評価= 適切	A-1	プレ自己評価とピア評価の根拠で不一致だった箇所に焦点化したやりとりの結果、本人は自己評価を適切に修正することができた	4	○	手順通りに進行
	A-2	プレ自己評価の根拠を投稿していない。ピアからの評価を見て、自己評価の誤りにすぐに気付いた	4	△	手順から一部逸脱
	A-3	プレ自己評価を開示しておらず、不一致に着目したやりとりはしていない。とにかくピアの意見を取り入れた	1	△	手順から逸脱
プレ自己 評価= 適切・ ピアから の評価= 不適切	B-1	プレ自己評価の根拠をピアに伝えたところ、ピアが評価を改めたので、本人は自己評価の得点と根拠を変更しなかった	1	○	手順通りに進行
	B-2	本人がプレ自己評価の根拠を投稿しなかったため、ピア評価が誤っている可能性を両者で検討する機会のないまま無批判にピア評価を受容した	3	×	手順から逸脱
	B-3	Google フォームで教員に提出したプレ自己評価とは異なる自己評価を協議ワークでピアに伝えたため、協議ワーク上はプレ自己評価とピア評価が一致していることになり不一致に着目したやりとりが生じなかった	3	×	手順から逸脱
	B-4	ピアからの評価の得点は不適切だが、コメントは妥当であり本人のプレ自己評価・根拠と一致。よって本人は自己評価の得点と根拠を変更しなかった	3	△	ピアが評価基準の理解不足
	B-5	自己評価の根拠を投稿しているが、レベル 4 を満たしているかどうか、ループバックに基づいた検討をしなかった。本人はピア評価を受容してしまった	2	×	両者が評価基準の理解不足

■学生 142 の協議ワークのやりとり (A-1 に分類) 教員評価 4、プレ自己評価 2、ピアからの評価 4、ポスト自己評価 3
ピア「(略) 構造とコミュニケーションを用いて適切に説明できていると思いました。(略) それに加え、「若者や風俗店・遊技場は社会的秩序を乱す」とみなされている暗黙了解について大学の論文など比較的信頼性のある文献を参考にして補足されている点は、一般的な報道のレベルを超えたものであると私は判断しました。よってこのレポートをレベル 4 の評価とします。」
本人「(略) 論理展開を裏付けるデータや論文がなく、評価基準表のレベル 3 の「客観性のために、事例・データやほかの文献を効果的に用いている」という条件を満たしていないのではないか? と思い、レベル 2 と判断しました。自分自身、まだ「客観性」という言葉の立ち位置についての理解があいまいなので、該当箇所の論理展開は客観的と言えるかどうか、〈ピア〉さんのご意見を伺いたいです。」^①

ピア「なるほど、そこまで考えが及んでいませんでした。〈学生 142〉さんのレポートを読ませていただいたとき、(略、……という論理展開は) 主観的だという印象を私は受けませんでしたが、〈学生 142〉さんの自己評価を読んで考え直してみました。私は十分客観的なレポートだと感じましたが、少しでも客観性を高めるには(……)」以下、略)

■学生 224 の協議ワークのやりとり (B-2 に分類) 教員評価 3、プレ自己評価 3、ピアからの評価 2、ポスト自己評価 2
Google フォームで提出した自己評価 3 の根拠: 「コミュニケーション→構造→コミュニケーションの説明が論理的に矛盾なく新聞記事を用いることで効果的に説明することができた^②と思うから。しかしながら、独自性に富むわけではなく一般論に留まるように感じるためレベル 4 には達していないと思う。」

ピア「レベル 2。全体的に「コミュ→構造」をもう少し具体的に説明して欲しいです。どのようなキーワードなのか、何に対する危機感なのかわかりづらいと思います。(略) その間の構造とコミュの関係性はごちゃごちゃしていて一読ではわかりづらかったです。特に「同調圧力や相互監視という構造を形成するとともに大多数の自粛(コミュニケーション)をもたらし、さらに大多数の自粛(コミュニケーション)は共通意識(構造)を強化した」の部分です。「相互監視と同調圧力(構造)」は位置付けや客観的根拠がよくわかりません。」

本人「丁寧に評価していただきありがとうございます。自分はレベル 3 であると評価していたのですが、もう一度読み返してみると確かにごちゃごちゃとしていてわかりにくいと感じ、評価 2 というのも納得しました。^③要素を盛り込みすぎたことで複雑になってしまい、根拠が不十分になってしまったように思います。論の流れがスムーズになり、それぞれの根拠が明白になるように書き直そうと思います。ありがとうございました。」

■学生 118 の協議ワークのやりとり (B-4 に分類) 教員評価 2、プレ自己評価 2、ピアからの評価 3、ポスト自己評価 2
ピア「私ならあなたのレポートを、レベル 3 と評価します。根拠ですが、コミュニケーションと構造の関係はしっかり説明できていますし、最初の新聞記事から更に深い内容になっているとは思いますが、「松下幸之助による日本人特性を参考にする」部分が、信憑性に薄くなってしまうのではないかと思います。^④私の考え方がずれているのかもしれませんが、松下幸之助の考えはあくまで松下幸之助という個人の考えでしかなく、なんの統計や調査でもないもので、根拠に乏しいものになっているとおもいます (略)」

本人「私自身も客観的という概念でこのレポートを書けているのかと確信が持てていませんでした。そこで、自分のレポート評価をレベル 2 としていました。松下幸之助を歴史上の人物と認識し、採用できるのではないかと考えました。しかし、レポートにおいて統計や調査が重視されるため、レポートの根拠として信ぴょう性の高いものを探し、後半部分を修正していきたいと思います。」(略)

図 4 LMS のフォーラムでの学生同士の協議ワークのやりとり (文意を損ねない範囲で誤字脱字を修正)

4の下線③からは、「ごちゃごちゃしてわかりにくい」というピアからの厳しい言葉を見た時点で自信がなくなってしまう、プレ自己評価が誤りであると判断してしまったと考えられる。

B-3の3名は、プレ自己評価としてGoogleフォームに提出したのとは異なる得点を協議ワークでピアに伝えてしまう、初歩的な手順の逸脱をしている。自分が提出したプレ自己評価を失念したとしても自動返信メールで確認できるはずだが、異なる得点を投稿してしまった。このようなことが生じた理由の1つとしては、他者から評価を受けるということは情意面に関連した複雑な認知活動であることを踏まえると、ピア評価を見た時点で自信がなくなってしまう、自分がつけたプレ自己評価の記憶をゆがめてしまった可能性が考えられる。

B-4は、ピアはレポートを正しく認識してはいるものの、ループリックを参照して得点に変換して本人に伝える際に、誤った得点を伝えてしまっている。学生118のピアが根拠として記述している内容は教員側から見て概ね妥当である（図4の下線④）。このようにレポートを捉えたならばループリックの得点は2になるはずだが、ピアは3と伝えている。根拠は概ね適切であったので、本人はピアの得点は受容せず、適切を維持できたという点では協議ワークは一部機能したといえよう。このプロセスが生じた理由としては、ピアが評価基準の理解が不十分であったことが考えられる。学生118のピアの場合は、ループリックのレベル2と3の差異が、コミュニケーションが構造に与える影響の説明が適切かどうかであることの理解が欠けていたといえる。このようなレベル間の差異を理解する目的で、各レベルに対応する4種のサンプルレポートを用いた評価練習を第5週に実施したが、この実施方法では不十分であった可能性がある。

B-5は、レベル3と4の不一致の理由を検討する際に、両者の評価基準の理解が不十分であるため、ピアの評価が正しいと誤解してしまっている。このプロセスが生じた理由もB-4同様、本実践の評価練習の実施方法では、評価基準の理解が不十分であった可能性が考えられる。

(3) 小括

以上の協議ワークのプロセスについて表4にまとめた。協議ワークの列の「○」は協議ワークが機能した、「△」は部分的にしか機能しなかった、「×」はほとんど機能しなかったと考えられることを意味する。「○」は21名中5名と少なかったことから、今回の「協議ワークを取り入れたピアレビュー」の実施方法には課題があるといえる。具体的には、協議ワークの手順から逸脱していたり、評価基準を一部誤解したままやりとりをすすめていたりする学生は、自己評価が不適切になるという課題があったといえる。一方で、プレ自己評価の根拠を投稿するなどの手順通りに取り組んだ学生は自己評価をより適切に修正することができたり、適

切さを維持できたりしていた。その結果として、分析Iの示すように受講生全体の平均としては自己評価力の向上がみられないという結果に繋がったと考えられる。

5. まとめと今後の課題

本稿では学生の意見交換を重視したオンデマンド型授業において、「協議ワークを取り入れたピアレビュー」をLMS上で実施し、効果とプロセスを検証した。その結果、受講生の平均的な傾向として「協議ワークを取り入れたピアレビュー」の前後で自己評価力の向上は、統計的に有意な差はみられなかった。自己評価力の向上はその後の期末レポートの改善に繋がっていることが示唆された。協議ワークのプロセスからは、プレ自己評価の根拠を書きこむという手順から逸脱している学生や、評価基準の理解が不十分な学生のペアにおいて、自己評価が不適切になっているという課題の存在が示唆された。

プレ自己評価の根拠を投稿するという手順を逸脱してしまった理由の1つとして、オンデマンド型授業で「協議ワークを取り入れたピアレビュー」を実施するにあたり、自己評価とピア評価の独立性を担保できるようにしたことと副作用として、手順が複雑になってしまったことが考えられる。

よりよい実施方法に向けて「プレ自己評価の根拠を投稿することの支援」「評価基準の理解の支援」の2点が必要と考えられる。

1点目の支援について具体的には、プレ自己評価の得点だけでなく根拠も投稿することがこの学習課題にとって重要であることを課題指示文で学生に伝えることで、学生がその意義を認識しながら取り組むことができる可能性がある。また、LMSのフォーラムを利用するのではなく、プレ自己評価の得点や根拠などワークに必要な事項を必ず入力する入力項目を設定したピアレビュー用のシステムを開発・利用することも解決になりうる。

2点目の支援については評価練習の実施方法について実証的な研究が今後の課題となる。本稿では評価練習の効果や、「協議ワークを取り入れたピアレビュー」への影響について、十分に検討することができなかった。前年度は「教室対面・グループ学習・異なる4テーマで4水準のサンプル」という条件であった（岩田, 2020a）のに対して本実践では「LMS・個人学習・同一テーマで4水準のサンプル」と様々な条件が異なっていたため、比較が困難であった。今後、条件を整えた実証的研究が必要である。

これらの改善策により協議ワークが教室対面型同様に効果を発揮すれば、より質の高い期末レポートを作成ことができ、それを通して知識を活用する力をより高めることができる可能性がある。

本研究の限界の1点目は、協議ワークのプロセスを検討する際に、ルーブリック得点のみに着目した点である。分析を通して明らかになったように、協議ワークの鍵を握るのは、ルーブリック得点だけでなくその根拠が重要だ。今回の分析では、得点は同じで根拠が異なるペアのやりとりを分析することはできなかった。

本研究の限界の2点目として、教養科目の社会学におけるレポート作成課題という文脈において得られた知見であるため、他の文脈的条件の授業でも同様の結果が生じるとは限らない点である。

本研究の限界の3点目は、本研究で論じた対象が知識活用力の育成を学習目標とするオンデマンド型授業であったが、知識活用力の育成を検証することはリサーチデザインの対象外とした点である。授業実践のデータから、学生の自己評価力の向上がレポートの改善に繋がる傾向は確認できたが、レポートの改善が即座に社会学的思考法を活用する力の育成を意味するわけではない。このような知識活用力の育成を検証するためには、評価課題の妥当性を確認したり、学習場面とは異なる設定の課題を用いて測定したりする必要がある。

今回のオンデマンド型授業では、学生らが指示した手順通りに進めることが予想以上に難しいことがわかった。学生らの反応を窺いながら学習活動の手順を説明できる教室対面型と異なる難しさが存在する。本稿は単一フィールドにおける事例研究であるが、より一般的な原則が導かれるための知見の蓄積に貢献できれば幸いである。

謝辞

コロナ禍という未曾有の事態の中、オンデマンド型授業に真摯に取り組み、さらに本研究にご協力くださった受講生の皆さまに深く敬意と御礼を申し上げます。

引用文献

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum.
- Earl, L. M. (2013). *Assessment as learning: using classroom assessment to maximize student learning (2nd ed.)*.

Thousand Oaks, CA: Corwin Press.

- Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal of Educational Statistics*, 6, 107-128.
- 岩田貴帆・田口真奈 (2020). 「パフォーマンス課題における自己評価力を高めるための協議ワークを取り入れた相互評価活動の開発」『日本教育工学会論文誌』43(Suppl.), 173-176.
- 岩田貴帆 (2020a). 「協議ワークを取り入れたピアレビューによる学生の自己評価力向上の効果検証」『大学教育学会誌』42(1), 115-124.
- 岩田貴帆 (2020b). 「協議ワーク後の自己評価の適切さがパフォーマンスの改善に与える影響」『大学教育学会第42回大会発表要旨集録』, 170-171.
- 松下佳代 (2016). 「アクティブラーニングをどう評価するか」松下佳代・石井英真 (編)『アクティブラーニングの評価』東信堂, 3-25.
- 小野和宏・斎藤有吾・松下佳代 (2018). 「PBLを評価する改良版トリプルジャンプにおける『学習としての評価』の要因」『京都大学高等教育研究』24, 35-44.
- Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education*, 28(2), 147-164.
- 佐渡島沙織・富永敦子・太田裕子 (2012). 「早稲田大学における学術的文章作成授業の成果：大学院生が個別指導するe-ラーニング初年次授業」『大学教育学会誌』34(1), 119-126.
- 佐渡島紗織・宇都伸之・坂本麻裕子 (2015). 「初年次アカデミック・ライティング授業の効果：早稲田大学商学部における調査」『大学教育学会誌』37(2), 154-161.
- Yucel, R., Bird, F. L., Young, J., & Blanksby, T. (2014). The road to self-assessment: exemplar marking before peer review develops first-year students' capacity to judge the quality of a scientific report. *Assessment & Evaluation in Higher Education*, 39(8), 971-986.

Effectiveness and Process of Peer Review with Discussion for the Moderation of Self- and Peer-Assessment for Enhancing Students' Self-Assessment Ability in On-demand Classes

Takaho Iwata¹, and Haruka Shibata²

(¹Graduate School of Education, Kyoto University, ²Graduate School of Human and Environmental Studies, Kyoto University)

Our research aims to propose how to implement “peer review with discussion for the moderation of self- and peer-assessment” in on-demand classes more effectively through three types of analyses. The data was collected from the peer review in a sociology course at Kyoto University. First, our findings imply that the average self-assessment ability of the students was not enhanced. Second, it was confirmed that the improvement of self-assessment ability leads to the improvement of students' performance. Third, the students whose self-assessment ability declined during the peer review either violated our instruction to send the evidence of self-assessment to the peer review or did not fully understand the assessment criteria. Therefore, instructing students on the importance of sending self-assessment or using the system to facilitate the procedure might improve the effectiveness of the peer review.

Keywords: Peer review, Self-assessment ability, Report assignments, On-demand classes