

AO・推薦入試を見据えたパフォーマンス評価の実証的研究

—社会科目のパフォーマンス課題の実践とその評価方法の検証—

中切正人（福井大学），雨森聡（静岡大学），大久保貢（福井大学）

本研究は、社会科目のパフォーマンス課題とその評価を国立大学の AO・推薦入試に活用して、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「大戦直前のドイツの中学校のカリキュラムと授業」を開発し、それに取り組む高校生の活動をルーブリックで評価した。活動場面ごとに測定される能力を思考力とコミュニケーション力に絞り、改訂版タキノミー等を参照して二つの能力の具体的構成要素を明示した。評価を担当した教員により、パフォーマンス課題とその評価方法・ルーブリックが評価された。その結果、情報過多と時間不足で測定不能な要素がある等の改良点が指摘され、次の実践への足がかりとなった。

1 本研究の目的と背景

1.1 本研究の目的

本研究は、国立大学における AO・推薦入試を念頭に実践研究したものである。2014 年の中央教育審議会の答申「新しい時代に(後略)」以降、高大接続改革の動きが加速された。高校教育現場への影響力が大きい入学試験において、東大・京大をはじめ AO・推薦入試の多面的・総合的な選抜方法が立ち上がった。そこで測定される能力は 21 世にふさわしい能力で、センター試験では測定が難しい能力である。

しかし、センター試験と異なり、AO・推薦入試で測定される能力の詳細・測定方法・評価方法の大半はブラックボックスの中にあるため、高校側は十分な対応ができていない。これでは、新しい時代にふさわしい高大接続改革が有効に機能しない可能性が高い。

そこで本研究は、AO・推薦入試での活用を前提に社会科目にかかわるパフォーマンス課題を開発し、そこで測定される能力の構成要素と測定方法・評価方法を具体的に明示する。そして、高校生を対象にパフォーマンス課題を実施し、評価者として参加した高大の教員からパフォーマンス課題とパフォーマンス評価に対する提言を得て問題点を整理する。以上の活動を通して、今後の AO・推薦入試における汎用的な選抜・評価方法の確立の基盤を構築することを目的とする。

1.2 本研究の背景

1999 年の中央教育審議会答申「初等中等教育と高等教育との接続の改善について」で大学全入時代に備えた多面的・総合的な評価が示されて以降、国立大学側では京都工芸繊維大学のダビンチ入試や福井大学の

大久保(2008, 他)をはじめ、理系分野を中心に AO・推薦入試の開発や実践研究が積み重ねられてきた。文系分野では、2014 年の中央教育審議会答申以降、有力大学でも AO・推薦入試の多様化が加速した。

これに対し高校側は、センター試験対策に追われて AO・推薦入試対策に二の足を踏み、そこで測定される多様な能力の育成に対応しきれていない。その理由の第 1 は、国立大学の募集人員の大半がセンター試験を伴う一般入試であるからである。第 2 に、一部の難関大学を除き、一般入試におけるセンター試験の比重が各大学の個別試験よりも高いからである。高校側から見ると、国立大学志望者の大半はセンター試験の得点により志願先の大学・学部の選定が左右される（谷口・山口編, 2011）。以上の 2 つの観点から、多くの進学校では受験生に対し、秋以降は年明けのセンター試験対策を核とした受験勉強への集中を望む。なぜなら高校側は、この間に AO・推薦入試志願者のセンター試験対策が遅れたり（志願者が不合格の場合）、担任教師の AO・推薦入試志願者への対応が本人や在籍クラスのセンター試験への集中力を阻害したりする可能性を避けたいからである（谷口・山口編, 2011）。

第 3 の理由は、AO・推薦入試で測定される能力は多肢選択式のセンター試験で測定される知識・技能に比べて、その構成要素などが不明瞭な高次の思考力・判断力・表現力等の能力だからである。それに加えて、大学・学部ごとにその試験内容・試験方法・測定能力・評価方法等の多様性が高く、それらの多くがブラックボックスの中にあるからである。たとえばお茶の水女子大学文系学部の新フンボルト入試においても、測定される能力の構成要素や採点方法・採点基準は非

公開である。AO・推薦入試のブラックボックスの推測は、予備校の把握する情報か、過去に受験者がいた場合にそこから得られる一部の情報に依存せざるを得ない。こうして高校側は大学・学部ごとに、かつ暗中模索的にブラックボックスに対応せざるを得ない。

これがセンター試験であれば、多肢選択式試験で測定される知識・技能は、思考力・判断力・表現力等に比べて明晰性や妥当性・信頼性が相対的に高い。その上、高校側は完全公開されたセンター試験の問題と解答を容易に入手し、過去問を十分に検討できる。その結果、出題される知識や技能を予測して受験生に対応することが可能である。以上より、高校側は大量の生徒を画一的に、能率よく指導することができる。これは裏を返すと、高校側はセンター試験で測定される知識・技能の修得を中核とした教育を行っていることになる（谷口・山口編, 2011）。

こうした中で本研究は、大学側が取り組む多面的・総合的評価の拡大に対する高校側の受け入れ態勢の整備を支援する観点から、上記第 3 の観念の克服を目指す。すなわち、高校側の国立大学 AO・推薦入試対策の活性化に貢献して、高大接続に寄与することを意図する。そのためには、大半の国立大学の AO・推薦入試で不明瞭な以下の 4 点を明確にし、高校側の対応を容易にすることが必要である。①試験や面接等で出題・質問され測定される能力とその構成要素は何か。②各能力を測定する出題・質問の具体例は何か。③出題・質問に対する解答・回答例は何か。④解答・回答に対する具体的な採点方法・採点基準は何か。本研究は公開の可能性を実証的に検証する第一段階である。

1.3 本研究の研究手法

本研究は、AO・推薦入試という限られた時間と空間をパフォーマンス課題の場に設定し、そこに資料分析に基づいて互いに討論する大学のゼミナール教育（次節で後述）を適用して、そこで発揮される高次の能力の構成要素を明示し、その評価方法を検証する。

すなわち、現行の AO・推薦入試で不明瞭な 4 点（前項で示した①～④）にそれぞれ対応する以下の(1)～(4)の研究手法をとる。(1)パフォーマンス課題（ゼミナール教育）として出題・質問され、測定される能力を思考力とコミュニケーション力に絞り、その構造と構成要素を示す。(2)それらの構成要素を測定する場面をパフォーマンス課題の遂行場面に位置づけ、課題の出題事例を提示する。(3)出題事例に対する回答例を、ループリックにおいて、各構成要素の活用・測定場面に求められる評価規準として示す。(4)回答

に対する採点方法・採点基準をループリックに示す。

さらに上記の(1)～(4)の研究手法は、パフォーマンス評価を担当する高校と大学の教員によって評価される。そして、パフォーマンス課題とループリックによる評価方法が批評され、問題点や課題が指摘される。

2 パフォーマンス課題と測定する能力の構成要素

2.1 パフォーマンス課題の開発ヴィジョン

2.1.1 フンボルト理念とゼミナール教育

近代大学を創設し「研究を通じた教育」を特徴としたフンボルト理念を推進する仕組みは、文系ではゼミナール教育（理系では実験室教育）である。そこでは学生がテーマについて自分で資料を調べ、その結果を学生・教員の前で報告し、互いに討論を交わしながら知識を確定していく。そして、その過程を通して未知の知識を自力で獲得し、発見の仕方を学び取ることが目指された（潮木, 2004）。

本研究は、この汎用性の高いゼミナール教育を高校と大学の文系学部（人文・社会科学）との接続の場に割り当てる。これにより、本研究の成果を一部の大学に留めることなく他大学との間で共有する可能性が生まれる。その結果、試験内容・試験方法・測定能力・評価方法等の多様性が高い AO・推薦入試を、ある程度収束させることが期待されると考えるからである。

2.1.2 ゼミナール教育とパフォーマンス課題

具体的に、本研究は社会科目のパフォーマンス課題の実践の場を、高校生が参加するゼミナール（グループワーク）とする。ここでは、高校生がパフォーマンス課題のテーマ「大戦直前のドイツのカリキュラムと授業」について自分たちで資料を検討し、互いに討論を交わしながら当時の国民中学校の時間割と授業内容を確定していく。その過程を通して高校生たちは思考力とコミュニケーション力を発揮し、未知の知識を獲得し、発見の仕方を学び取る。そして、グループワークの成果を参加者全員の前でプレゼンテーションする。

この一連の活動がパフォーマンス評価の対象とされる。したがって、本研究で評価される対象は高校生のゼミナール活動であり、ここで高い評価を得た生徒は大学入学後の活躍が期待される（本研究のパフォーマンス評価の妥当性と信頼性が高いと仮定した場合）。

2.2 パフォーマンス課題のシナリオと課題文

パフォーマンス課題の開発にあたり「逆向き設計」論を参照した（西岡・田中編, 2009）。

まず、パフォーマンス課題の中核にかかわる 2 つの「本質的な問い」と、それに対応する「永続的理

解」を設定した。すなわち、(1)「権力はどのように自己システムを安定させるか?」に対し、「コントロールされた教育の利用」。(2)「ナチズムはどのように政権維持し国民を戦争に駆り立てたか?」に対し、「次代を担う子供たちの教育を対象に歴史や科学を歪曲したナチスの世界観を与え、戦争を正当化」。

次に、上記の本質的な問いに対してそれを問わざるを得ないようなパフォーマンス課題を考案するにあたり、以下の6要素「GRASPS」を設定した。

パフォーマンス課題の目的(Goal)はナチズムの教育利用と教育統制の方法・内容の理解であり、生徒が担うシミュレーションの役割(Role)はナチスの教育関係者の立場(カリキュラムの作成者)である。次に、その相手(Audience)は、教育対象の子供・生徒であり、そこに想定されている状況(Situation)は、ナチス政権獲得後から第二次大戦(1933-39)までの国民中学校である。そして、生み出すべき作品・実演(Product, Performance)は国民中学校のカリキュラムと授業内容であり、その評価の観点(Standards and criteria for success)は、思考力とコミュニケーション力である。

以上の作業を経て、以下の6つのパフォーマンス課題(A)~(F)を設定した。それは、(A)ナチスの時代(1930年代後半)の中学校の「男子のカリキュラム」を考える。(B)「女子のカリキュラム」を考える。(C)「体育」・(D)「歴史」・(E)「理科」の授業内容を考える。その上で、(F)ナチスの時代の「カリキュラム全体の特色」を整理・分析・考察する、である。

2.3 パフォーマンス課題で測定する能力の構成要素

2.3.1 思考力の構成要素と設定の根拠

思考力の構造と構成要素の設定にはアンダーソンらの改訂版ブルーム・タキソノミーを参照して、以下の①から⑥に向かってより高次になる6個の構成要素を設定した。本研究のパフォーマンス課題は社会科目の「世界史」にかかわる課題であることから、これらの構成要素は歴史的思考力と深くかかわるものである。

①想起力：史実(事実的知識)を認識し想起する能力。
 ②理解力：歴史概念(概念的知識)を解釈し理解する能力。
 ③適用力：歴史的資料(映像と文献)から史実や歴史概念を批判的に導き出す能力。いわゆる史料批判。
 ④分析力：ナチズムの教育システムと第二次大戦(戦争)の歴史的因果関係を分析する能力。
 ⑤評価力：ナチズムの教育システムのプラス面とマイナス面を両面から評価する能力。

⑥創造力：ナチズムの時代と現代の教育システムを比較・考察することによって、歴史を学ぶ意義を構想したり、新たな歴史認識を作り上げたりする能力。

なお、これらの構成要素と学校教育法第30条に示された学力の三要素との対応関係は次のようであると考える。すなわち、「基礎的・基本的な知識・技能」は①想起力、②理解力、③適用力に該当し、それを活用して課題を解決するために必要な「思考力・判断力」は④分析力、⑤評価力、⑥創造力に該当する。

2.3.2 コミュニケーション力の構成要素と設定の根拠

コミュニケーション力の構成要素は、先行研究において様々な見解が見られるため、いくつかの文献を参照した上で(末田・福田, 2003; ほか)、報告者が独自に以下の4点を設定した。

- ①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。
- ②伝達工夫力：他者の意見に対し論理的な質問を発したり、自分の考えを伝達する工夫をしたりする能力。
- ③共同創作力：他者と共同して一つの研究成果(作品)をまとめ、完成させる能力。
- ④共同表現力：他者と共同して一つの研究成果(作品)を公の場で分かりやすく表現する能力。

なお、これらの構成要素は学校教育法第30条に示された「表現力」に該当するものとする。

表1 パフォーマンス課題の進行と評価項目の対応関係

パフォーマンス課題の進行		思考力	コミュニケーション力
1 日 目	13:30 - 15:00	・映像視聴と講義(90分) ・確認テスト(10分) = 知識として測定	想起力 理解力
	15:15 - 17:00	グループワークA: ・男女別カリキュラム、歴史・理科・体育の授業内容の検討・考察	適用力 分析力① 伝達工夫力 共同創作力
	9:10 - 10:00	グループワークB: ・前日の成果の報告 ・新グループとしてカリキュラムと授業内容の整理・分析・考察	適用力 分析力② 伝達工夫力 分析力③ 評価力
2 日 目	10:00 - 10:45	グループワークC: ・研究成果のまとめ ・発表用資料の作成	評価力 伝達工夫力 創造力① 共同創作力 創造力②
	11:00 - 11:50	プレゼンテーション(個人として分担発表) ・男カリ・女カリ・体育・歴史・理科の授業内容	伝達工夫力 共同表現力① 共同表現力②

3 パフォーマンス課題と評価

2016年12月17・18日に実施したパフォーマンス

課題とその評価の概要を【表 1】に示す。

3.1 参加者の募集と事前準備

パフォーマンス課題に取り組む高校生と評価担当の教員（以後「評価者」と表記）の募集対象は、これまで福井大学と高大連携あるいは「福井ラウンドテーブル」で関わりの深い学校とした。その結果、高校生 32 名「1 年生 10 名（県内 8，県外 2），2 年生 22 名（県内 13，県外 9）」と、評価者 10 名「高校側 8 名，大学側 2 名」の参加を得ることができた。

参加に応募した高校生には、実施日の 1 か月前にプリント資料を送付した。これは高校生たちが実施日のパフォーマンス課題を考察する際の基礎知識となり、第 1 日目に視聴する映像資料の理解を深めるプリント資料（映像の概要とナレーション部分の記述，市販の世界史資料集から現代史関係頁の抜粋）である。

3.2 パフォーマンス課題第 1 日目の実践と評価

3.2.1 想起力と理解力の測定

第 1 日目は、最初に「NHK スペシャル 映像の世紀 第 4 集 ヒトラーの野望（75 分）」を視聴し（報告者による解説付き），パフォーマンス課題の下準備を整えた。

そして、視聴直後に紙ベースの「確認小テスト：多肢選択式全 14 問（10 分）【図 1】」を用いて、想起力と理解力を測定した。たとえば、想起力はゲルニカの作家という「いつ・どこで・誰が」レベルの史実（事実的知識）の測定であり、理解力は世界大恐慌の影響という歴史概念（概念的知識）の測定である。

2. スペイン内戦の最中（中略）「ゲルニカ」を描いた画家を一人選べ。<正解(2)「想起力」>
 (1) マティス (2) ピカソ (3) ゴッホ (4) ルノアール

3. ヒトラーが政権を握った 1933 年前後の世界について誤りのある文章を一つ選べ。<正解(2)「理解力」>
 (1) 1929 年の世界大恐慌により、経済不況が世界中に広がり、アメリカ国内では労働者がデモを繰り返し、農民は生活苦にあえいだ。
 (2) 独自の社会主義経済路線を歩んでいたソビエト連邦でも、世界大恐慌の影響を受けて国内は混乱し、強制労働が行われた。 (3) 以下、略

図 1 確認小テスト「想起力と理解力の測定」抜粋

3.2.2 思考力とコミュニケーション力の測定

映像視聴後、「グループワーク A（100 分）」では最初に高校生のグループ分けが行われ、前節で掲げた(A)～(E)のパフォーマンス課題文を検討する 5 つのグループ（5～6 人）が結成された。そして、5 つのグループにそれぞれ一つずつ課題文が割り当てられた。たとえば、課題文(A)の男子のカリキュラムを作成するグループでは、事前配布資料と当日の映像資料を参

考にしなが、新たに示されたわが国の中学校 2 年相当のカリキュラムを記したプリント資料を参照し、用意された時間割ボード【図 2】に科目を配置する作業を行った。そして、グループのメンバー同士で意見を交換し合いながら時間割編成作業を進めた。時間割編成は、あらかじめ用意された科目カードを時間割ボードに貼り付ける形で進められた。その際報告者から、新科目を追加することが可能であり、各科目の時間数にはその重みとしての意味があることが解説された。

評価者は、1 グループに 2 名が割り当てられ、各個人の活動の様子をルーブリックに従って評価した（ルーブリックの一部を【表 2】に示す）。この場面では、思考力では「適用力」と「分析力①」、コミュニケーション力では「相互理解力」と「伝達工夫力」と「共同創出力」が評価された。

3.2.3 「適用力：史料批判」測定の意義

第 1 日目の作業の中核は事前配布資料と当日配布資料の適用であり、それを評価する適用力の活用場面は、翌日のグループワーク B での活用場面と合わせて本研究の中核をなす。その位置づけは、ベルリン大学創設期の「古典学ゼミナール」が膨大な文献資料に取り巻かれた専用の部屋で行われたことに由来する（潮木，2004）。これは、お茶の水女子大学の新フンボルト入試（文系）の場が図書館であるのと軌を一にする。さらに、理系のフンボルト理念の推進装置が実験室教育の実験であることと好対照をなす。なお、適用力の測定は近代歴史学の基盤かつ歴史教育においても重要な「史料批判」に重なると思われる。

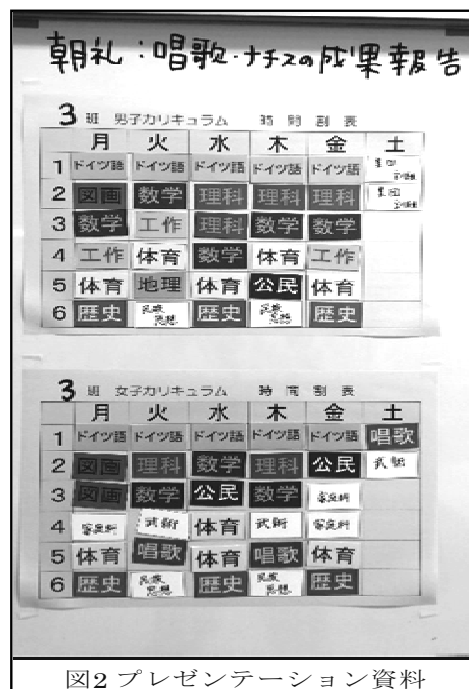


図 2 プレゼンテーション資料

3.3 パフォーマンス課題第 2 日目の実践と評価

3.3.1 グループワーク B で考察の最終場面

第 2 日目の最初は、前日のグループワーク A の成

果を活用してパフォーマンス課題(A)～(E)を再分析・考察した上で、(F)カリキュラム全体の特色を考察するグループワーク B である。前日の 5 グループが解体され、各グループから 1 人ずつ集められ、5 人 1 セットの新グループが結成された。最初に前日のグループワーク A の成果「男女別カリキュラム、体育・歴史・理科の授業内容」が互いに報告され、新メンバーでそれらを共有し合った。そして、新グループとして、グループワーク B の課題に取り組んだ。

この評価場面では、思考力では適用力と分析力に加えて評価力が測定され、コミュニケーション力では相互理解力と伝達工夫力が測定された。

3.3.2 グループワーク C で研究成果のまとめ

グループワーク C では、新グループで検討・考察してきた研究内容を整理し、プレゼンテーションする準備作業に入った。男女別カリキュラムは最終的な時間割がボードに配置された。体育・歴史・理科の授業内容は口頭で発表する準備が進められ、報告者から「なぜそのように授業内容を考えたか？」を重点的にまとめ、発表することが指示された。

この評価場面では思考力では評価力に加えて新たに創造力が測定され、コミュニケーション力では伝達工夫力と共同創造力が測定された。

3.3.3 グループワーク D のプレゼンテーション

最後のグループワーク D は研究成果発表会である。どのグループも 5 つの発表内容「男女別カリキュラム、体育・歴史・理科の授業内容」は充実していた。

たとえば、【図 2】は第 3 班の男女別カリキュラムである。報告者から指示されていなかった「朝礼」まで配慮されている。なぜ、ドイツ語を朝の 1 時間目に毎日配置したか、あるいは、土曜日の「集団訓練」の時間はどのようなもので、なぜ位置づけたか等の説明は、どれも説得力に富む内容であった。

この場面では、コミュニケーション力で伝達工夫力に加えて、新たに共同表現力が測定された。

3.4 修了証明書の送付と参加生徒の感想

後日、参加生徒へは評価者による評価データを基に作成された「修了証明書」が送付された。これは、参加者の進学先等への有効活用を意図したものである。

表 2 パフォーマンス課題のルーブリック(一部抜粋)

測定する能力の構成要素、および、その評価規準	評価基準(評点)			
	1	2	3	4
適用力 ・グループワークにおいて、事前に提示された資料、あるいは生徒がこれまで積み重ねてきた学習・体験等に基づく何らかの資料を、歴史的資料として認識した上で引用し、自分の考え等を解説、検証、導き出すことができる能力。	・事前に提示された資料も、生徒のこれまで積み重ねも引用されない。	・事前に提示された資料が、生徒のこれまでの積み重ねが、どちらか一つを引用した上で、自分の考えや歴史的な見方を解説、検証、導き出す。	・事前に提示された資料と、生徒のこれまでの積み重ねの両方を引用した上で、自分の考えや歴史的な見方を解説、検証、導き出す。	・事前に提示された資料と、生徒のこれまで積み重ねを複数引用した上で、自分の考えや歴史的な見方を解説、検証、導き出す。その際、引用資料には史料批判(立場、宗教、男女等)が加えられている。
分析力 ③ ・ナチズムの世界観(反ユダヤ主義等)や、第二次大戦の構成要素(民族固有の領土、資本主義対共産主義等)が、それぞれどのように関係(強弱、+、-、包含、対立、等)しあっているのか識別、構造化(組織化)する能力。		・ナチズムの世界観や第二次世界大戦の構成要素との関係性を想定することができる。	・ナチズムの<中略>関係性のうち、二つのペア、もしくは三つの構成要素の関係性を想定することができる。	・ナチズムの<中略>の関係性のうち、三つ以上のペア、もしくは四つ以上の構成要素の関係性を想定することができる。
評価力 ・各グループで分析された男女別カリキュラムや歴史・理科・体育の授業内容について、そのプラス面(知育偏重ではない体育重視等)とマイナス面(戦争賛美等)の両面を確認するとともに、批評することができる能力。		・ナチズムの教育システムのプラス面とマイナス面の両面を確認、批評することができる。	・ナチズムの<中略>の両面について、ともに一つずつ確認、批評することができる。	・ナチズムの<中略>の両面について、ともに二つ以上確認、批評することができる。
創造力 ① ・ナチズムの教育システムと今日のわが国の教育システムを関連付けることにより、各構成要素を新しい教育システムに再組織化し、未来の教育に新たな仮説を設定することができる能力。	・事前に提示された資料と授業日の映像やグループワークを通して得られた情報の認識・理解に留まり、そこから新たに「わが国の将来の教育のあり方」や「政治と教育のかかわり方」や「男女の教育の将来的展望」等々の発展的課題(テーマ)を論じたり、設計したり、構想したりすることができる。	・事前に<中略>に留まることなく、そこから新たに「わが国の将来の教育のあり方」や「政治と教育のかかわり方」や「男女の教育の将来的展望」等々の発展的課題(テーマ)を一つ提示することができる。	・事前に<中略>の発展的課題(テーマ)を一つ提示することができる。そして、その輪郭を論じたり、設計したり、構想したりすることができる。	・事前に<中略>の発展的課題(テーマ)を二つ以上提示することができる。そして、それぞれのテーマについてその輪郭を論じたり、設計したり、構想したりすることができる。
相互理解力 ・他者の話す内容に対してメモを取るなどして向き合い、その意見に合理的な理解・判断を下したり、共感したりすることにより、グループで議論しやすい場を作り出すことができる能力。	・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。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、何らかの質問を返す反応が見られない)。	・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、もしくは内容に対する何らかの質問が見られる)。	・他者の<中略>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、もしくは内容に沿った的確な質問が見られる)。	・他者の<中略>見られる)。さらに、同意や質問を通して、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。
共同創造力 ・グループワークにおいて、他者と共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で、以下の4つの観点から構成されている。1.グループのメンバーに敬意を払い、礼儀正しく振舞っている。2.明確な話し方や表情や身振り手振りなどを交えて、自分の前向きな姿勢を伝えている。3.メンバーの能力を信頼し、メンバーがそれぞれの課題を十分に達成できるという確信を伝えることで、メンバーのやる気を高めている。4.メンバーを手助けしたり、励ましたりしている。		・左に示した4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4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4つの観点のうち、すべての観点を行うことができる。
共同表現力 ① ・プレゼンテーション時において、発表者がメンバーと共同しながら、話術(姿勢、ジェスチャー、アイコンタクト、声の表現)を駆使できる能力。	・グループの他のメンバーからの支援があるのに、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。	・グループの他のメンバーからの支援を受けて、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。	・グループの他のメンバーからの支援を受けて、発表者の話術がプレゼンテーションに興味深くしており、発表者が落ち着いて見える。	・グループの他のメンバーからの支援を受けて、発表者の話術がプレゼンテーションを積極的に示しており、発表者が洗練され、自信があるように見える。

パフォーマンス課題に参加した生徒の感想には、自己を見つめなおし、その成長を評価する記述が多かった。たとえば、「他校の生徒とディスカッションすることにより、自分自身や学校を客観的に眺める機会になった」、「1 日目と 2 日目でグループのメンバーが異なったので、いろいろな考え方を聞くことができ、自分の意見を深く考え直した」である。

4 まとめと今後の課題

4.1 評価に対する評価者の提言と評点の分析

パフォーマンス課題について、評価者から以下のような提言があった。まず、実質 1 日の日程の中で、歴史的資料が多すぎて情報過多となり、情報の整理が不十分、もしくは討論に深みが欠けた。その結果、新しい科目などの創意工夫を試みる余裕が無くなり、評価力や創造力を発揮する余裕が無かった。

次に、パフォーマンス評価では、高次の思考力の活用場面をどのグループワークの中で行うかが問題となった。初日からほとんど発言しなかった生徒が二日目のグループワーク C の段階で創造力を発揮したケースに対し、それまでの過程の評価は十分できたのであろうかという指摘である。また、コミュニケーション力の測定では、休憩時間の活動が重要な活用場面として利用されるべきであると指摘された。他にも、事前にルーブリックを提示する方法や、グループワークを数台のカメラで撮影し、後に映像を見ながら個人の行動評価を測定する方法が指摘された。

表3 能力別評点の占有率と一致率(%)と平均値および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	重み付 きκ係 数**
	1	2	3	4	0差	1差	2差	0と1		
適用力	5	35	48	12	42	51	7	92	2.7	0.91
分析力	11	38	32	19	36	39	25	75	2.6	0.85
評価力	22	60	10	8	35	46	19	80	2.0	0.86
創造力	47	26	18	9	8	38	54	46	1.9	0.72
相互理解力	15	27	29	29	38	36	27	74	2.7	0.84
伝達工夫力	7	36	33	24	38	46	16	84	2.7	0.88
共同創作力	18	30	26	25	36	37	27	73	2.6	0.84
共同表現力	10	29	47	14	35	50	16	85	2.7	0.88
全体	15	35	32	18	33	43	24	76	2.5	0.85

*評価場面ごとに班編成や評価者が異なるが、総じて全体の7割近くの生徒が二人の評価者に評価されている。 **κ係数は2次の重みの計算式による。

【表 3】は、能力の構成要素別に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、さらに重み付きκ係数を示したものである。一人の参加者に対する二人の評価者の評点差を見ると、どの構成要素も完全一致する「0 差」は 50%を超えていないため評価者間信頼性は高くない。ただし、創造

力を除いた項目では、「0 と 1」差が 8 割程度見られ、重み付きκ係数の値も 0.81 を超す高い一致率となっている。今後改良する際の参考にしたい。

次に妥当性について見ると、評価力と創造力の評点の平均値は、他の構成要素の平均値が 2.6 もしくは 2.7 であるのに対して極めて低く(2.0 と 1.9)、評点の度数分布も下位に偏っている。評価者の提言にもあるように、時間的余裕の欠如から参加者がこれらの能力を発揮できなかったと考えられ、他の構成要素に比べて評価項目の妥当性が低いと考えられる。他の構成要素の妥当性については、今後、同様のパフォーマンス評価を積み重ねて検証していきたい。

4.2 まとめと今後の課題

本研究では、国立大学文系学部の AO・推薦入試での活用を前提に、社会科目のパフォーマンス課題を開発して実施し、評価者によりルーブリックに基づいて思考力とコミュニケーション力が測定された。その結果、評価者からパフォーマンス課題のあり方やその測定方法・評価方法・評価基準など多くの指摘を得た。

次の実践では、情報過多に陥らないようにパフォーマンス課題の内容を修正する必要がある。また、パフォーマンス評価については、評価場面やルーブリックを再検討し、評価項目・評価基準・評価尺度の妥当性を高め、実践を積み重ねることにより信頼性を高める必要がある。今後は、評価・測定・社会科教育等々の専門家のご指導を仰ぎながら、改良に努める必要がある。その上で、AO・推薦入試に活用される汎用的な選抜・評価方法の基盤作りに挑戦していきたい。

参考文献

- Anderson, L. W., et al (Eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing*, Longman.
- 西岡加名恵・田中耕治編 (2009). 『「活用する力」を育てる授業と評価・中学校』学事出版。
お茶の水女子大学ホームページ
<<http://www.ocha.ac.jp/news/h280126.html>>
(2017年2月18日)
- 大久保貢 (2008). 「高大連携活動 (SPP など) に参加した入学生の入試成績と初年度成績」『大学入試研究ジャーナル』, 18, 25-30.
- 末田清子・福田浩子 (2003). 『コミュニケーション学——その展望と視点』松柏社。
- 谷口典雄・山口和孝編 (2011). 『センター試験—その学力に未来はあるか—』群青社。
- 潮木守一 (2004). 『世界の大学危機』中央公論社。