

Current Situations and Issues on English Proficiency Assessment in Classrooms : A task Analysis Designed by High School Teachers in Shizuoka

メタデータ	言語: jpn 出版者: 公開日: 2019-03-28 キーワード (Ja): キーワード (En): 作成者: 出口マクドナルド, 友香理, 福田, 純也, 亘理, 陽一 メールアドレス: 所属:
URL	https://doi.org/10.14945/00026365

高等学校における英語運用能力アセスメントの現状と課題

静岡県立公立高校のパフォーマンス・タスク分析

出口マクドナルド 友香理・福田 純也・亘理 陽一

(静岡大学研究補佐員・元静岡大学特任助教・静岡大学)

Current Situations and Issues on English Proficiency Assessment in Classrooms:

A Task Analysis Designed by High School Teachers in Shizuoka

Yukari Deguchi McDonald, Junya Fukuta, and Yoichi Watari

Abstract

The purpose of this study is to clarify the current situations and issues of English proficiency assessment in high schools through analyzing performance tasks designed by high school teachers. Ninety-two performance tasks from eighty-six public high schools in Shizuoka were examined. The analysis was conducted to see if the tasks teachers made were accurate at the A2 CEFR level (Common European Framework of Reference for Languages) and to discover any general tendencies or trends amongst Japanese high school teachers in task design. The CEFR descriptor, and the reference for teacher training created by the Comprehensive Education Center of Shizuoka Prefecture were both applied for the analysis. The results showed that only 38 tasks were considered to meet the criteria of CEFR level A2, which reveals a gap between the actual criteria and the teachers' perception of the task level. Moreover, the variation of writing and speaking tasks were found to be quite limited. As one prominent characteristic, almost all the writing tasks were argumentative, and the rubrics utilized to assess students' performance were unclear. Distinguishing the differences in each evaluation scale within the rubrics were inconsistent with the contents of the tasks. With regard to speaking tasks, the majority were presentation tasks and the number of interactive tasks were nearly half that of presentation tasks, which could possibly be caused by the teachers' unwillingness or hesitation to create interactive tasks. From the results, it can be assumed that a lack of experience high school teachers have in creating performance tasks has led to discrepancies in task variation and understandings of CEFR levels. In conclusion to correct these issues within schools, it is necessary to develop and provide teacher training programs to gain knowledge and experience in implementing performance assessment.

キーワード：高校英語、パフォーマンス・タスク課題、タスク分析、CEFR

1. はじめに

高等学校の英語教育はかつてない変革を迫られている。2018年度告示の外国語科学習指導要領では、「外国語を使って何ができるようになるか」という観点から「ヨーロッパ言語共通参照枠」(Common European Framework of Reference for Languages, CEFR)を参考にした技能別の指標形式の教育目標案(CAN-DO リスト)が示され、小中学校と接続において、2020年度の指導要領全面实施時に60%以上の生徒が高等学校卒業時にA2～B1レベルの英語運用能力に達することを目指すとしている(文部科学省, 2015)。一方、同省の2017年度の英語力調査調査事業によれば、A2レベル以上を有する生徒の割合は「聞くこと」33.6%、「話すこと」12.9%、「読むこと」33.5%、「書くこと」19.7%であり、いずれの技能においても2013～2017年度の目標50%には達していないのが現状である(文部科学省, 2018a)。

同時に、高大接続改革の一環として2020年度から導入される大学入学共通テストにおいて、英語に関してのみ4技能評価に、「大学入試英語成績提供システ

ム」の参加要件を満たすと確認された民間の英語資格・検定試験(認定試験)を活用する方針が示されている(文部科学省, 2018b)。他方で、その受験料負担や各地域で受験可能な試験の偏り、試験会場・監督・採点等の実施体制、性格の異なる複数の試験間の比較妥当性・基準の信頼性、学習指導要領との整合性などの観点から問題点や懸念が多く指摘されている(南風原(編), 2018)。

こうした状況から認定試験の扱いについて大学ごとの判断が揺れる中で、東京大学は、2020年度実施の一般入試については、認定試験の成績以外に、「CEFRのA2レベル以上に相当する英語力があると認められることが明記されている調査書等、高等学校による証明書類」、もしくはそのどちらも提出できない事情を明記した理由書のいずれかを出願の要件とした(東京大学, 2018)。上述の問題点・懸念から認定試験受験を必須としない判断には理があると言えるものの、他方で、高等学校側に英語力の証明を委ねることも信頼性や実行可能性の点で問題があり、高等学校の外国語科教員に責任が投げられる事態を招き

かねない状況となっている。

静岡県においては、「各中・高等学校の外国語教育における『CAN-DO リスト』の形での学習到達目標設定のための手引き」（文部科学省，2013）を受け、県内の全ての公立高等学校が CAN-DO リストを作成しており、静岡県総合教育センターの指導の下、その運用・改善の取り組みを求められている。同センターは、指導と評価の一体化の方針のもと、CAN-DO リストに対応する目標課題を単元に位置づけたアセスメント・サイクルの実施を進める取り組みの一環として、昨年度末に各高校からライティング・スピーキングのパフォーマンス課題と具体的なパフォーマンス事例を収集し、筆者らにその分析を依頼した。

本論は、この分析を通じて、公立高等学校における英語運用能力アセスメントの現状と課題を明らかにしようとするものである。つまり、生徒が A2 レベルの英語運用能力を有しているかどうかを見極める課題を作成したと外国語科教員が言うとき、本当にそれは A2 レベルを測る言語タスクとなっているかということであり、生徒のパフォーマンスに対して妥当な英語運用能力の評価を下すのに十分なアセスメント能力を教員が持っているかどうか問われることになる。この検討は、上述の高校英語教育の変革への対応としてのみならず、養成課程・研修における外国語科教員の力量形成においても極めて重要な意義を持っている。なぜなら教員のアセスメント能力の涵養こそが、学習者の多様性や英語運用能力の実態に応じた的確な指導を行う上で要だからである。

実際に作成・実施されたパフォーマンス・タスクを言語評価論の観点から分析する前に、次節では、英語教育におけるパフォーマンス評価の現状を整理すべく、心理学的・教育学的評価論の理論的成果を概観し、パフォーマンス評価実施における課題を述べる。

2. 英語教育におけるパフォーマンス評価

2.1 パフォーマンス評価とは

Kane, Crooks, & Cohen (1997)は、テストや評価にはある意味必ずパフォーマンスを伴うことから、Performance Assessment（パフォーマンス評価）という言葉に修飾語として Performance が重ねられている意義を見出そうとした。Performance には、評価したいものを直接的に観察する方法と、間接的に解釈されたアウトプットを観察する方法があることを主張する。例えば、どの程度風景が描けるかを評価するために、実際にキャンパスに描くパフォーマンスを求めるのが直接的評価方法であり、風景画を描くための知識や理論をレポートで報告するパフォーマンスを求めるのが間接的評価方法である。どちらもパフォーマンスを伴うテストまたは評価方法であり、間接的評価方法で観察できる知識や理論は風景画を描く上で関連している。

しかし、実際にどの程度風景画を描くことができるかを判断するには直接的評価を行う必要がある。この区別から Kane, Crooks, & Cohen (1997)はパフォーマンス評価の特徴を、評価するものと実際に観察するものが非常に類似している、または、近接していることだと捉えている。言い換えれば、可視できるパフォーマンス全てが必ずしも評価すべきものを測っているとは限らないという、評価目的に対するパフォーマンス評価方法妥当性の重要性を説いている。（第二言語を含む）言語学習の観点からパフォーマンス評価を解釈する Yu (2014)は、「パフォーマンス評価」(Performance Assessment)は、たった一つの正解を選ぶこと以上の力が必要となる評価であれば該当すると解釈できるような広い意味と曖昧さをもつ言葉になっていると主張する。その一方で、パフォーマンス評価では課題（タスク）の解決が必須となることから、“Task-Based Performance Assessment”という言葉でしばしば置き換えられる。加えて、課題（タスク）デザインの際に実世界で起こり得る課題に類似するような状況設定や、リアルであるやりとりを行うという「真正性」が重要な要因となっていることも、状況的真正性と相互作用性真正性のどちらをより重要視するかという議論はあるにせよ、パフォーマンス評価を特徴づける点だと言える。

言語運用能力に対するパフォーマンス評価は、他教科におけるそれとは性格が異なり、言語自体が評価の構成概念であると同時にパフォーマンスの手段そのものとなる。例えば、理科で酸化物を還元するというパフォーマンス評価では、評価手段として、酸化銅を還元して銅を取り出す方法をレポートにまとめるという言語表現と、実際に実験を行うことが手段になり得る。しかし、言語運用能力のパフォーマンスでは、アウトプットの手段として、言語を用いる以外に他の手段が存在しないため、常に言語が評価するためのツールとしての役割を担うだけでなく、その言語化されたアウトプット自体が評価内容となる。そのため、評価したい能力の構成概念や課題の設定、評価基準の作成、評価結果の解釈が特に難しくなり、評価する言語が外国語になればその複雑さはさらに高まることが予想できる(Yu, 2014)。それを踏まえて外国語の言語運用能力評価を具体的に考える際、言語諸能力(Language Abilities)とタスク達成(Task Completion)という評価判断の対象について、大きく2つの大きな考え方が存在する。上記でパフォーマンス評価と Task-Based Performance Assessment が互いに置き換えられる場合があると発言したが、実は両者はその視点が大きく異なっていることに注意しなければならない。

パフォーマンス評価では評価対象を言語諸能力とし、その中の評価したい要素をパフォーマンスの中で引き出すようにタスクがデザインされるため、タスク達成

の有無は評価に必ずしも必要ではない。外国語を用いて表現したいことを言語化するプロセスでは母語での処理に比べ、より多くの言語的知識・技能が必要となることから、文法や特定の表現を使用できるのかを評価対象とすべきという考え方である(Brown, 2004)。一方、Task-Based Assessment はあくまで学習者が与えられた課題をどの程度達成できたかが評価の対象であり(Long & Norris, 2001)、課題達成のために使用する言語材料を前もって特定し評価・分析をすることはないという考え方に立っている(松村, 2012)。Long (2015)は、第二言語を使用する状況においてタスク達成の他にも、言語運用能力の要素として文法的、社会言語学的、語用論的、文化的に言語を適切に使用することが求められる状況があることを認めつつも、こうした要素を評価する難しさと、そうした状況が求められる頻度の低さを指摘し、言語運用能力の評価においては第一に意味交渉によるタスク達成を評価の対象とすることを主張している。

言語運用能力を評価するために、両立場の観点が存在することから、実際に言語運用能力を評価するパフォーマンステストを行う際には、評価の対象をタスク達成か、ある言語諸能力であるかとのどちらか一方に決めるのではなく、評価配分に差はあるが両方を取り入れたものが多い(Yu, 2014)。評価者は、言語を用いて情報を受け取ったり発信したりする能力を評価する上で、学習者のニーズや何を身に付けさせたいのかを明確化した上で、そのバランスの按配を考えることが求められる。

2.2 中学校・高等学校外国語(英語)教育におけるパフォーマンス評価の位置づけ

教育評価方法は国や地域、その時代の教育観・学力観によって大きく変化する。日本の学校教育の中でパフォーマンス評価という言葉が出現した背景には、2000年以降 OECD が実施する国際的学習到達調査の一つである PISA 調査の大きな影響がある。PISA 調査では、知識を単に覚えているのではなく、現実世界や実生活の中で知識・技能を活用したり応用したりすることができる力が現代の学力の一つとして捉えられ、活用・思考・判断を必要とする問題が調査に導入されている(松下, 2010)。これがパフォーマンス評価であり、2010年には「パフォーマンス評価」という文言が文部科学省発行の文書「児童生徒の学習評価の在り方について(報告)」の中で登場し、これを思考力・判断力・表現力の評価において活用することが奨励されることとなった(文部科学省, 2010)。西岡・石井・田中(編)(2015)はパフォーマンス評価を、実世界の中で起こり得るような場面の中で、思考力・判断力・表現力・創造性を身につけさせようとする「真正の評価」(Authentic Assessment)の立場から、「知

識やスキルを使いこなす(活用・応用・総合する)ことを求める問題や課題などへの取り組みを通して評価する評価方法の総称である」(p. 10)と定義している。

外国語(英語)教育に焦点化すると、パフォーマンス評価は、到達度評価としての特性をもつことから、「言語を用いて〇〇することができる」という視点で作成されている CAN-DO リストとの親和性が非常に高い。実際に 2011 年には、文部科学省による「国際共通語としての英語力向上のための 5 つの提言と具体的施策」の中で、学習指導要領に基づき生徒が到達すべき英語力を身につけるとともに、その評価を教員の指導改善に活用することを目的として、英語を使用して具体的に何ができるかという行動目標の形で学習到達目標を設定することが提言された(文部科学省, 2011)。その中でパフォーマンス評価の活用が奨励され、その後、本提言が具現化される中で各中・高校において「CAN-DO リスト」の作成が促されることとなった。2013 年 3 月には、文部科学省による「各中・高等学校の外国語教育における『CAN-DO リスト』の形での学習到達目標設定のための手引き」が作成され、CAN-DO リストの形での学習到達目標設定の目的は、4 つの評価観点のうち「②外国語の表現の能力」および「③外国語理解の能力に関して生徒が身に付ける能力」を各学校が明確にし、主に教員が生徒の指導と評価の改善に活用することと示されている(文部科学省, 2013)。

学校教育で目指す学力・学習の質は、知識の獲得と定着の段階である「知っている・できる」、知識の意味理解と洗練段階である「わかる」、知識の有意な使用と創造ができる段階である「使える」の 3 段階が存在し、そのうち、パフォーマンス評価は「わかる」と「使える」段階の学力評価を網羅していることから(田中, 2011)、CAN-DO リストを用いて生徒の英語での表現力および理解力が身につけているかを評価するためには、パフォーマンス評価の中でも高次の学力である「使える」レベルまでを評価することが必要であると解釈できる。使えるレベルの評価には、様々なパフォーマンス評価の中でも、「知識や技能を応用できるような実生活の中で起こり得るような文脈を与えて、あるふるまいを行わせ、それをもとに評価を行う」パフォーマンス課題での評価が求められる(西岡・石井・田中(編), 2015, p. 135)。京都府立園部高等学校の実践から例を借りれば、「学校の中で、あなたのお気に入りの場所や場面はどこ(いつ)ですか? お気に入りの場所(場面)を写真にとって、どんな風に素敵なのか、なぜ好きなのか、クラスの人に語ってください」という課題や、「ALT の先生に対して、まず自分自身のことを少なくとも 5 つの英文を使って述べ、その後それについての英語による質問に答えながら対話を続けてください」、あるいは「もうすぐ卒

業ですね。これまで生きてきた 18 年間で振り返って英語で自分史を書いてください」といった課題が挙げられる(田中, 2016, p. 93)。リーディングとリスニングの中で言語メッセージを受け取り解釈するのみのタスク設定では、学習者の言語知識・技能の活用や応用のプロセスが見えないため「わかる」という段階の学力を評価することは可能であっても「使える」レベルの英語での表現力を評価することは極めて難しい。同時に、英語による表出を伴うタスクといえ、スピーキングとライティングの活動全てが表現力を評価するものとはなり得ない。単にある単語を発音することやある文法を使用し例文を書くというテストは知識・理解の評価に留まっている。「使える」レベルに達するための段階の一つとして、知識や技能を評価することは大切であるにせよ、それだけで、知識の有意味な使用と創造ができる段階である「使える」レベルの表現力が身につけているかどうかを評価することは出来ない。日常生活の中で「使える」かどうかは絶えず試されるわけではない現状において、学校教育の一環としての英語教育においては、CAN-DO リスト形式に示される、生徒に身につけてほしい英語の 4 技能が、「知っている」・「わかる」レベルの評価に留まらないよう、「使える」レベルの評価が可能であるパフォーマンス課題の作成および実施が求められているのである。

2.3 パフォーマンス評価実施における課題

中学校・高等学校における英語科のパフォーマンス評価の実践・取り組みから、実施する上でパフォーマンス課題(タスク)およびルーブリックの作成が大きな課題の一つであることが窺える(愛知県立惟信高等学校, 2014; 日本英語検定協会, 2015; 田中, 2016)。パフォーマンス課題の作成では、まず評価内容を評価できる課題設定になっているかという妥当性の問題があり、評価の目的と手段が一致するデザインが行わなければならない。「使える」レベルを評価するためのパフォーマンス課題には多様なものがあり、エッセイ、小論文、論説文、物語の作成といった作品として評価するものや、プレゼンテーション、ディベート、演劇のように動きとして可視できる「実演」をもって評価されるものがある。タスクを行う際には、種類、形式、難易度、評価方法といった要因が変化することで学習者のパフォーマンスの結果に影響を与えることがあるため、同様なタイプに偏ったタスク設定や一回の総合評価のみで学習者の言語力の評価することは避けるべきである(Long, 2015; 松村(編), 2017)。また、単元の最後に単にスピーキングやライティングタスクを行うという単純で断片的なものではなく、評価時期や目的に合う教材や単元でタスクデザインされるべきであり、3 年間の学びを見通して体系的にパフォーマンス評価・課題を計画する必要がある(田中, 2016)。

評価基準作成においても、その基準が評価対象となる要因を評価できるものになっているかという妥当性の問題に加えて、信頼性を高めることが求められる。答えが一つで正解か不正解かを安易に評価することが出来ないパフォーマンス評価では、評定スケールとその内容を記述したルーブリックが用いられる。評価対象者一人に対して同様のルーブリックを用いた複数の評価者間での評価のゆれだけでなく、複数の評価対象者を同一の評価者が同じルーブリックを使用したとしても評価のゆれが起こることが予想される。評価の信頼性を高めるために、複数人によるルーブリックの作成・実践を行うことが奨励されるだけでなく、個々の評価者である教員のパフォーマンス評価をはじめとする質的評価を行うための目を養うことが求められている(Yu, 2014)。またルーブリックは、ある特定のタスクや単元レベルの短期的なものから、学期、学年、校種の区切りといような長期的なものが存在するが、評価結果が最終ゴールではなく子どもたちが現在の学習到達度を把握し、ステップアップするためには何が必要か教師と子どもが共に共通認識されるためのものとして、学習者(評価を受ける者)、教師(評価を行う者)に分かり易く示されるものでなくてはならない(田中, 2011)。

3. 高校英語授業のパフォーマンス・タスク分析

高等学校における英語授業のパフォーマンス評価はどのような実態にあるだろうか。平成 29 年度に静岡県内の公立高等学校を対象として静岡県総合教育センターより依頼を行い、「CEFR A2 レベルのスピーキング(インタラクション・プレゼンテーション)またはライティング力を測定するために用いたパフォーマンステスト」の作成・実施が行われた。

「評価の目的と手段が一致するかどうか」という意味での教員のパフォーマンス・タスク作成能力を検討すべく、以下では、依頼された通りの「CEFR A2 レベルのタスクになっているか」という点や、教員がどのようなタスクを作る傾向があるかについて、言語評価論の観点から分析を行う。CEFR 分類の基準に関しては、CEFR が示す能力記述子(ディスクリプタ)(Council of Europe, 2001)、およびそれをもとに静岡県総合教育センターが教員研修において使用したタスク自己評価の基準を用いた。

分析の対象となったのは、静岡県内の公立高等学校 86 校の 92 のタスクである。以下の分析では、回収した課題の中で、分析者が内容を把握できたもののみを扱い、分析不可能であったものを除いている。対象学年や評価方法等、静岡県総合教育センターが依頼した記述項目に記載がないものもあったため、合計値が必ずしも校数や総タスク数に一致しないことを付記しておく。

まず、すべてのタスクをレベル別に分類した結果は下表の通りである。要求通りの「A2」レベルであると判断されたタスクの数は 38 となり、これは全体の 46% 程度である。レベル別の分布を見てわかるように、対象となった教員は全体的にタスクの難易度を低めにして作成する傾向があるといえる。

レベル	数
A1	31
A2	38
B1	13
B2	1

作成されたタスクの内、教員が「A2 レベルのタスク」と認識しているものの対象学年をみると、1 年生が 24 (28%)、2 年生が 35 (41%)、3 年生が 26 (19%) という内訳になっている。

この数値は、生徒が A2 レベルの課題を達成すると教員が見込んでいる学年とも見ることができる。高校生活前半を対象とする課題が全体の過半数を超えているところを見ると、前述した、教員がタスクの難易度を低めに作成するという傾向が、「教員が生徒の習熟度が低いとみているため」であるとはいえないことがわかる。

次に、作成されたタスクがインタラクティブ、プレゼンテーション、ライティングのいずれを対象としているかをコーディングし、その内訳を計算した(カテゴリーが複数にまたがるものもある)。結果として、ライティングの総数が 51 (54%)、インタラクティブが 15 (16%)、プレゼンテーションが 29 (31%) という結果となった。このことから、パフォーマンステストを行う場合、対象教員の作成するタスクの過半数がライティングとなるという結果が示された。今回回収されたタスクが、必ずしも日常的に教員が作成するタスクそのものを代表しているとは限らないが、この結果は、インタラクティブタスクを作ることに教員の抵抗感を反映している可能性がある。

次に、作成されたタスクを質的に検討した。まずライティングタスクの際立った特徴として、ほぼすべての課題が「○○(すべき)である、賛成か反対か」といった *argumentative* な産出を求めるものであった。ライティングにはさまざまなジャンルがあるが、教員の作成する課題のジャンルがごく限定された範囲にとどまることが示唆される。また、生徒が例文をそのまま使うことが可能なものや、1 つのタスクで例文の下線部だけ変えて書くといった形式のものも散見された。このような課題が、前述のように「知っている」・「わかる」レベルの評価に留まらない「使える」レベルの評価が可能であるパフォーマンス課題であるといえるかは疑問である。

またライティング課題の評価に関しては、ルーブリックがわかりにくいものが多くみられた。その具体的内容としては、記述が不明瞭でどのように評価すべきかわからないもの(例えば、1 点を与える基準を示す記述 A と 2 点を与える記述 B の違いが不明瞭なもの)、作文内容と評価の関連が全く見られないもの、内容がいかなるものであっても一定の語数を切ると大幅に減点されるもの(かつなぜその語数が必要であるか不明な目的設定のもの)が指摘できる。後者の二つの特徴からは、達成される課題の目的や到達目標を明確にし、それにむけて授業を行うという、授業内容と評価内容の一致が行われていない可能性が示唆される。

次に、プロダクション課題を検討したところ、こちらにはほぼすべての課題が「書いたものを覚えて発表する」という形式をとるものであることが目立った。この形式を持つことが即問題であるということにはならないが、これもスピーチのもつ様々な目的や形式にかかわらず、極めて限定された課題しか行われていないことが示唆される。

インタラクティブ課題については、そのほとんどが、あるものを示しながら説明を行う「Show and Tell」の形式をとっていた。また、「スキットの内容を入れ替えて演じる」という課題もみられた。このことは前述の、課題形式の限定性が指摘できるだけでなく、教員はインタラクティブ課題を作ることに困難を示しているという仮説とも合致する。

上記の特徴とは異なるインタラクティブ課題はごく少数しか見つけることができなかったが、以下にそれら 3 つの課題を取り上げる。まず 1 つ目の課題は、授業で取り扱った 4 つの販売テクニック (*foot-in-the-door, door-in-the-face, low-ball, hard-to-get*) のうちの 1 つを選択し、その手法を駆使して教員に T シャツを販売する 1 分程度の口頭試験である。2 つ目の課題は、授業中に取り扱ったと思われる「Science of Love の 3 つの学説」を取り上げ、その学説をうまく説明して友達の悩みにアドバイスをするという課題である。そして 3 つ目の課題(これはプレゼンテーション課題として提出されたものであり、量的にもそのように扱っているが、その後発表したもののやり取りを含んでいるためここで紹介する)は、身の回りにあるユニバーサルデザインを持つものを探し、それについて、それが重要かどうか、将来広まると思うかなど、自分の意見を書き、発表を行い、それに対して意見交換をするものである。対象を見つけ出し、自分の意見を書き出したうえで、それを用いて発表・やり取りを行うという、言語使用に至るまでのプロセスを意識した課題である。

これら 3 つの課題に共通する特徴は、学習した学説や理論といった有意義な内容を取り上げ、実際にそれらを駆使して言語使用を行い、その言語使用を評価するというものである。これは、第 2 節でその重要性を

指摘した、「評価の目的と手段が一致するデザイン」となっているということが挙げられる。また言語使用の目的が明確であるため、生徒が学習した言語と使用する状況を結び付けて課題に臨むことができるといえる。このようにこれらの課題は、パフォーマンス課題を作成する際に重要と思われる観点を含んでおり、参考にすべきものであると考えられる。

一方タスクタイプにかかわらず全体的な結果としては、どのようなシチュエーションで行われる言語使用なのかかわからないものが3つのタスクでみられた。また、参照できる例文が非常に多い課題が散見された。参照できる例文が多すぎると、それらのパフォーマンス課題は、前述した例文の単語を入れ替えたスキットや、下線部を入れ替えるドリルと同様のものになってしまう。これらも「使える」レベルの評価が可能であるパフォーマンス課題であるといえるか疑問が残る。

4. 総合考察および今後の課題

教師が作成した課題を分析した結果としては、教師は課題の難易度を下げて作成する傾向にあることが明らかになった。ただし、それは教師が生徒の習熟度を低く見積もっているわけではなく、タスクを作成することに教師が不慣れであることが原因ではないかと考えられる。同様の理由によるものと予想される特徴として、教師が作成する課題のジャンルは非常に限定的であることも指摘された。この現状から予想される帰結は、1節で言及した大学入試に係る英語力の証明を高等学校側に委ねた際、実際はA2レベルより低いレベルのパフォーマンス課題しか経験していない生徒がそれ以上の運用能力を有すると判断されたり、ごく限られた範囲での英語使用経験しか持っておらず、大学入学後の幅広い要求には必ずしも応えられない可能性があるということである。言語使用はどのような状況で行われるかという点に関して教師は理解を深め、様々なジャンルでの言語使用への対応を学び、それを評価する必要があるだろう。

本稿でその重要性を述べたように、課題作成に関しては、その課題で用いられる言語使用の目的を明確にし、生徒が学習した言語と使用する状況を結び付けることによって、評価と目的を一致させる必要がある。また、ルーブリックの記述や評価方法にみられたいくつかの問題から、教師が生徒の産出を評価する際にどのような観点からどのように評価を行うかという点について、知識を深めることができるような機会を設ける必要があるだろう。教師による運用能力評価を妥当性・信頼性を有するものとするためには、単元を計画する際、タスク達成について明確で達成可能なゴールを設定し生徒と共有するとともに、その達成に必要な言語諸能力を具体的にした上で、生徒の産出を診断的・形成的・総括的に評価していくことが求められる。

教員がそのような評価能力を身につけていくための諸条件の考察や教員研修プログラムの開発は今後の課題としたい。

参考文献

- 愛知県立惟信高等学校 (2014). 「愛知県立惟信高等学校の取組 (外国語(英語)科) 高等学校における多様な学習成果の評価手法に関する調査研究」『愛知県総合教育センター平成 27 年度研究成果報告書』 23-46.
Retrieved from <http://www.apec.aichi-c.ed.jp/shoko/hyouka/hyoukashuhou/hyoukashuhou1.html>
- Brown, B. J. (2004). Performance assessment: Existing literature and directions for research. *Second Language Studies*, 22(2), 91-139.
- Council of Europe (2001). *Common European Framework of Reference for Language: Learning, Teaching, Assessment*. Cambridge University Press. Retrieved from <https://www.coe.int/en/web/common-european-framework-reference-languages>
- 南風原 朝和(編) (2018). 『検証 迷走する英語入試: スピーキング導入と民間委託』 東京: 岩波書店.
- Kane, M., Crooks, T., & Cohen, A. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18(2), 5-17.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*: Chichester: Wiley-Blackwell.
- Long, M. H., & Norris, J. M. (2001). Task-based language teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching* (pp. 597-603). London: Routledge.
- 松村昌紀 (2012). 『タスクを活用した英語授業のデザイン』 東京: 大修館書店.
- 松村昌紀(編) (2017). 『タスク・ベースの英語指導』 東京: 大修館書店.
- 松下佳代 (2010). 「PISA で教育の何が変わったか: 日本の場合」『教育テスト研究センターCRET シンポジウム報告書』 1-10.
- 文部科学省 (2010). 「児童生徒の学習評価の在り方について (報告)」中央審議会 Retrieved from http://www.mext.go.jp/b_menu/shingi/chukyo/chuky_o3/004/gaiyou/attach/1292216.htm
- 文部科学省 (2011). 「国際共通語としての英語力向上のための5つの提言と具体的施策」 初等中等教育局国際教育課外国語教育推進室 Retrieved from http://www.mext.go.jp/b_menu/shingi/chousa/shotou/082/houkoku/1308375.htm
- 文部科学省 (2013). 「各中・高等学校の外国語教育における『CAN-DO リスト』の形での学習到達目標設定のための手引き」 Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1332306.htm
- 文部科学省 (2015). 「生徒の英語力向上推進プラン」 Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikokugo/icsFiles/fieldfile/2015/07/21/1358906_01_1.pdf
- 文部科学省 (2018a). 「平成 29 年度英語教育改善のための英語力調査 事業報告」 Retrieved from http://www.mext.go.jp/a_menu/kokusai/gaikokugo/1403470.htm

- 文部科学省 (2018b). 「大学入学共通テストの枠組みで実施する民間の英語資格・検定試験について」 Retrieved from http://www.mext.go.jp/a_menu/koutou/koudai/detail/1408564.htm
- 日本英語検定協会 (2015). 『外国語教育における「CAN-DO リスト」の形での学習到達目標設定に関する現状調査』 Retrieved from http://www.shogaisoken.or.jp/htmltop/toppage.files/candochosa_2014.pdf
- 西岡加名恵・永井正人・前野正博・田中容子・京都府立園部高等学校・附属中学校 (編) (2017) 『パフォーマンス評価で生徒の「資質・能力」を育てる』 東京: 学事出版.
- 西岡加名恵・石井英真・田中耕治(編) (2015). 『新しい教育評価入門』 東京: 有斐閣.
- 田中耕治 (編) (2011). 『パフォーマンス評価: 思考力・判断力・表現力を育む授業作り』 東京: さんぎょう.
- 田中容子 (2016). 「高等学校におけるパフォーマンス評価の実践」 『カリキュラム研究』 25, 91-98.
- 東京大学 (2018). 「2020年度 (2021年度入学者選抜)以降における入学者選抜方法の検討について」 Retrieved from https://www.u-tokyo.ac.jp/ja/admissions/undergraduate/e01_admission_method_03.html
- Yu, G. (2014). Performance Assessment in the Classroom. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 615-630). London: John Wiley & Sons.