

ISSN 2187-4123

实用英語教育学会紀要

SPELT JOURNAL

第6号

*The Society of Practical English Language
Teaching*

2016 12

实用英語教育学会

目次

巻頭言	実用英語教育学会会長	鈞 晴彦 ……………	1
President's Message	The Society of Practical English Language Teaching, President	TSURI, Haruhiko	

【研究論文】

【実践研究】

「英語会話」授業におけるアウトプット評価方法としてのルーブリック 使用の成果と問題点について……………			3
--	--	--	---

Results and Issues of Using Rubrics for Output Tests in “English Conversation” Classes

北海道札幌英藍高等学校	宮越 哲史
Hokkaido Sapporo Eiai High School	MIYAKOSHI, Tetsushi

はじめてルーブリック評価を経験した日本人大学生英語学習者の評価 及びフィードバックに見るルーブリック評価実施上の留意点（2）……………			21
--	--	--	----

Some Issues in Using Rubrics in the Classroom: Assessment and Survey Results with Japanese
College Learners of English with No Previous Experience with Rubric Assessment, Part 2

北海道平取高等学校	渋谷 奈緒美
Hokkaido Biratori High School	SHIBUYA, Naomi

【投稿規程】……………			40
-------------	--	--	----

SPELT Journal 第6号の発刊に寄せて

実用英語教育学会 (The Society of Practical English Teaching) の紀要発行は、これで第6回目、通巻第6号となる。

今回、本紀要には2本の研究論文が寄稿された。

1本目のテーマは「『英語会話』授業におけるアウトプットテスト評価方法としてのルーブリック使用の成果と問題点について」、もう1本は、「はじめてルーブリック評価を経験した日本人大学生英語学習者の評価及びフィードバックに見るルーブリック評価実施上の留意点(2)」である。共に前回の論文をさらに進め、系統的に研究考証したものである。

この2本の論文は、共にルーブリック評価を授業で活用した実践研究論文である。前者は高校の「英語会話」の授業で、後者は大学の初年次の一般英語の授業でルーブリックを評価に活用した実践報告だが、どちらも統計的な手法を使ってデータをしっかり分析しているので、異校種の英語授業でルーブリック評価を行う場合でも参考になるだろう。1年から2年にわたる実践に基づくデータからは様々な情報が読み取れ、これからルーブリックを評価に取り入れようという者にとって大変有益な論文となっている。どちらの論文も、一昨年「学習を促進する評価のデザイン」というテーマで本学会が共催した評価に関するワークショップがきっかけになったと聞いているが、その意味で、ワークショップリーダーのカリフォルニア大学サンディエゴ校の當作靖彦教授が与えた影響は大きい。

今年2016年11月26日、その當作靖彦教授が再び北海道を訪れ、今度は札幌旭丘高校で「評価のパラダイムシフト—講義中心の教育からアクティブラーニング中心の教育への変化は評価をどのように変えたか」と題して講演を行い、その後、パフォーマンス評価、オーセンティック評価のアクティブラーニングの効果的な評価のデザインと実施方法としてワークショップを行った。前号の紀要でルーブリック評価の理論と実践を発表してくれた札幌旭丘高校の稲毛先生らを中心として企画された研修会であるが、本学会も協力団体として参画した。小中高大から、英語以外の複数教科にわたる教員が参加し、熱気に満ちた研修会であった。他教科の教員と評価の視点の情報交換をすることは、お互いに新たな気づきや発見に繋がっていく。内容を十分満足のいくところまで深める時間はなかったが、学校が組織的に教科を越えてルーブリック評価を採り入れていくための作業の方法論を学ぶ機会としては大変有意義な時間であった。

當作靖彦教授によれば、評価は、目的により伝統的評価と代替的評価をうまく組み合わせることで学習者をより適切に評価し、教育・学習効果を上げる必要があるという。また、学習が「アクティブラーニング」と呼べるものとなるための条件は、教員が明確な達成目標を学習者と共有すること、高次の思考を促す活動を使用すること、内容をその深層まで学習させること、また、教える内容と学習者の現実生活とのつながりに気づかせ、学習者中心、学習者主体で、共有した目標達成に向けて有意義な対話を学習者同士に行わせること、その中で、教師と学習者、学習者と学習者が社会構築的に互いに支援し合える環境を整え、同時に学習者自身がその学習

過程を内省できるように支援することだという。アクティブ・ラーニングは、現在様々な研修会が行われているが、どれにも共通する基本軸は、教員主体から学生主体となる学びへの転換であり、評価を上手に利用することで一層学習の効果を高めることに繋がるのだと力説されていた。英語で何ができるようになることが最終目標なのか、学期のはじめに、その目標の到達度を段階別に明示し、学期末に向けていま自分がどの段階に達しているかを自ら評価したり、互いに評価し合えるようにすることで、学習者が自分の足りないところを自省し、確認するルーブリック評価は、まさに、アクティブラーニングそのものであると言える。

今回寄稿された研究論文は、数こそ2本にとどまったが、共にデータの収集や分析に多くの時間を費やして完成に漕ぎ着けたもので、内容的にも、学校現場の実態を分析・考察した詳細な研究として貴重である。どの校種の教員にとっても大いに参考になり、良き指標となっていくものである。ルーブリック評価は、いま文科省が強く推進しているアクティブラーニングに直結している重要なテーマである。実用英語教育学会は、今後も多くの皆様のご意見をもっと頂きながら、このテーマのために真摯に研究を一層進めていく所存である。

実用英語教育学会会長
札幌学院大学
釣 晴彦

「英語会話」授業におけるアウトプット評価方法としての ルーブリック使用の成果と問題点について

宮越 哲史 (北海道札幌英藍高等学校)

Results and Issues of Using Rubrics for Output Tests in “English Conversation” Classes

Tetsushi MIYAKOSHI (Hokkaido Sapporo Eiai High School)

Abstract

Ample opportunities to perform output activities should be given to learners, especially to Japanese high school students, to develop their ability to communicate with others in English. According to the “Output Hypothesis,” optimal output activities can help learners develop their communication skills. Therefore, output tests such as presentations, speeches, and interview tests can help high school students develop the ability to speak. Teachers, however, must not forget to evaluate student output according to well-defined criteria. In addition, they should help students understand the criteria for evaluating student performance on output tests. Rubrics can clearly show the criteria to both teachers and their students. Using rubrics for evaluating student performance on output tests would help them understand the criteria and set targets. Therefore, teachers should use rubrics as an evaluation tool and monitor student progress regularly using rubrics. Regular monitoring of student progress using rubrics can help teachers and their ALT colleagues enhance the quality of their classroom instruction. Moreover, teachers should give their students the same rubrics they use, and regularly collect and collate the rubric scores they give to their students as well as the ones students give to themselves. Constant collating of data can provide teachers and their ALT colleagues with useful information about the effectiveness of their teaching methods. This paper mainly discusses the issues as well as the results of using rubrics for evaluating student performance on output tests in English class, based on investigations carried out in the author’s “English Conversation” classes.

1. はじめに

「英語会話」は、高等学校学習指導要領（平成 21 年 3 月告示）で示された外国語に属する科目のひとつである。筆者の勤務校では、「英語を通じて身近な話題について会話する能力を養うと同時に、海外での生活に必要な基本的な英語表現を使って会話する能力も養う。」という指導目標を設定し、Assistant Language Teacher（言語指導助手、以下 ALT）とのティーム・ティーチングでその授業を行うことを基本としている。また、複数年次（2・3 年次）が一緒に学ぶクラスと 3 年次単独のクラスがある。前者は 1 クラス、そして後者は 2 クラスあり、どのクラスも単位数及び授業内容は同じである。教科書は *My Passport*（文英堂）を使用している。また、評価方法としてルーブリックを活用し、3 種類のアウトプットテストを実施している。1 つ目は、2 人または 3 人のグループごとに作成したオリジナルダイアログを発表するプレゼンテーションテスト、2 つ目は、指定されたテーマに沿って自分の考えを発表するスピーチテスト、3 つ目は、2 人または 3 人のグループごとに生徒同士で指定された時間を有効に活用して英語で会話させるインタビューテストである。その得点を彼等の成績に反映させている。また、教師と生徒のルーブリック得点の一致度を確認して生徒がルーブリックの評価基準をどのくらい正確に把握しているのか確認する必要があると考えて、教師用のルーブリック評価表と同じものを生徒に配布して、アウトプットテスト実施時に自己採点をさせて回収している。

本論文では、「英語会話」の授業で取り組んでいるアウトプットテストの実践とその評価方法であるルーブリックの使用の成果と課題について報告したい。

2. 取り組みの契機

昨年度は、アウトプット活動を通じて、自分の発話を自分で検討できるだけでなく、インプットに対しても意識が高くなり、第二言語習得の様々な認知プロセスに大きな影響を与える可能性がある（鈴木・白畑、2013）という考えを参考にして、「アウトプット仮説」を教室現場に生かすことを試みると同時に、研究エッセイ、実験、写真の展示など違った種類のプロジェクトをクラスの生徒達が準備した場合に生じる評価の過程における主観性や公平性といった問題への答えをルーブリックが与えてくれる（ヤング & ウイルソン、2013）という考え方からヒントを得てアウトプットを重視した活動を授業に取り入れた場合の生徒の到達度を測る測定方法として評価の客観性と公平性が保証されるルーブリックの活用を試みた。また、生徒に授業の振り返りをさせる目的で授業アンケートを実施したが、アウトプットテストの自己評価をさせなかったために、教師評価と生徒自己評価の差を分析することができなかった。

今年度は、学習者に教師と同じルーブリックを用いて、別個に評価させ、それを比較できるようにする方法がある（當作・中野、2013）ことに注目し、教師評価と生徒自己評価の差を分析し、テスト結果に対する両者間の認識の差を確認して指導に生かしている。また、教師が評価を記入したルーブリックを生徒に手渡し、次の学習到達目標を設定させている。その取り組みについて、報告したい。

3. 実践の内容

3.1 調査対象とした「英語会話」クラス

本実践研究は、筆者が担当する3つの「英語会話」クラスの4つの生徒群を調査対象として行った。その生徒群とは、3年次の「英語会話」クラスの2生徒群と、2・3年次混合クラスの「英語会話」クラスの2つの生徒群の計4つで、各生徒群は以下のような構成である。

(1)	3年次生徒群1 (以下、3年(1))	3年次の複数クラスに所属する生徒8名	男子4名 女子4名
(2)	3年次生徒群2 (以下、3年(2))	3年次の複数クラスに所属する生徒10名	男子6名 女子4名
(3)	2・3年次混合3年生徒群 (以下、2・3年混合3年)	3年次の複数クラスに所属する生徒7名	男子3名 女子4名
(4)	2・3年次混合2年生徒群 (以下、2・3年混合2年)	2年次の複数クラスに所属する生徒8名	男子2名 女子6名

また、今年度8月22日～8月26日にかけて実施した授業アンケートに基づき、英語に自信がある生徒群（以下、英語得意自己認識群）と英語に自信がない生徒群（以下、英語不得意自己認識群）に分類して、英語に対する自信の度合いがアウトプットテストの取り組みに影響するかどうかの検証を試みた。

3.2 授業で行う3つのテスト

「英語会話」の指導目標の達成状況を測定する方法として活用しているプレゼンテーションテスト、スピーチテスト、及びインタビューテストの実施内容は以下の通りである。また、その3つのテストの教師得点は生徒の成績に反映される。

3.2.1 Presentation Test (プレゼンテーションテスト)

プレゼンテーションテストは、ルーブリック評価基準を生徒に周知させ、単元ごとに実施するテストであり、生徒が2～3名のグループに分かれて各単元で学習した重要表現を用いてオリジナルのダイアログを完成させて発表するテストである。ALTが出張などでその評価を担当できなくても定期的に実施することができるように日本人教師が評価を担当している。

今回研究対象とした単元は、レッスン1、2、及び3である。レッスン1は「留学先のホストファミリーのホームパーティー」、レッスン2は「ホテル宿泊」、そしてレッスン3は「観光案内所」の場面設定である。したがって、例えば、レッスン3であれば、観光案内所で必要な英語表現を学習するため、道案内をする時に必要な英語表現が重要表現となる。生徒はグループごとにその表現を用いてオリジナルのダイアログを作成して原稿を見ないで発表しなければならない。

3.2.2 Speech Test (スピーチテスト)

スピーチテストは、ルーブリック評価基準を生徒に周知させ、1ヶ月に1回の割合で5月から実施している。また、ひとつの授業内に4から5名の生徒がそのテストを受けている。評価者はプレゼンテーションテストと同様の理由により日本人教

師である。指定されたテーマに基づき、その原稿を彼等自身で作成して発表するテストである。例えば、生徒たちは「興味があることについて」というテーマで、その原稿を作成し、それを見ないで発表しなければならない。

3.2.3 Interview Test (インタビューテスト)

インタビューテストは、ルーブリック評価基準を生徒に周知させ、1ヶ月に1回の割合で5月から実施している。また、ひとつの授業内で全員がそのテストを受けることができるように配慮している。評価についてはALTの担当である。生徒はペアまたは3人グループで、制限時間を有効に活用して英語で会話しなければならない。そのテーマは、例えば連休の過ごし方や学校祭の印象などについてである。ペアに割り与えられる制限時間は2分間であり、3人グループに割り与えられる制限時間は2分30秒である。ペアまたは3人グループの会話終了後に生徒はALTの質問に答えなければならない。

3.3 授業で使用しているルーブリック評価表

上述の3つのテストのそれぞれについて、以下の表1、表2、表3のそれぞれに示すルーブリックを4月の最初の授業で配布して説明している。配布時に、それぞれの基準（各表の縦列）と水準（横列）が具体的に指し示す能力を、日本語を交えながら説明し、基準スコアを示すことで生徒に理解させている。また、教師が評価と得点を記入したルーブリックを生徒に手渡し、それを見て次の学習到達目標を設定させている。

(1) 表1 Presentation Test は、単元ごとに実施している「プレゼンテーションテスト」の評価表である。

表1 Presentation Test: Teacher: Student: Date:

	Excellent 10	Good 8	So-so 6	Poor 4
Grammar	There is no grammatical error.	There are a few grammatical errors.	There are some grammatical errors.	There are many grammatical errors.
Memorization	The student doesn't need any help and doesn't pause to remember.	The student doesn't need any help, but pauses to remember once or twice.	The student doesn't need any help, but pauses to remember several times.	The student needs help.
Pronunciation	The pronunciation is perfect.	The student makes one or two mistakes.	The student makes several mistakes.	The audience can't understand what the student is saying.
Voice	The student's voice is loud and very clear. The audience can hear easily.	The student's voice is clear and loud enough. The audience can hear everything.	The student's voice is quiet or unclear. The audience can't hear them sometimes.	The audience can't hear the student.
Tone	The tone changes in a natural way.	The tone changes.	The tone only changes a little, or sounds strange.	The tone is flat and does not change.
Expression	The student expresses emotion that feels natural.	The student expresses emotion.	There is a little bit of emotion.	There is no emotion.
Gesture	There are lots of gestures that look natural.	There are lots of gestures.	There are only some gestures, or they seem strange.	There is no gesture.

(2) 表2 Speech Test は、各単元のプレゼンテーションテスト実施日を含めた授業で一つの授業につき4名程度ずつ実施している「スピーチテスト」の評価表である。

表 2 Speech Test: Teacher: Student: Date:

	Excellent 10	Good 8	So-so 6	Poor 4
Grammar	There is no grammatical error. The expressions are easy to understand.	There are a few grammatical errors, but the expressions are understandable.	There are some grammatical errors. The expressions are not so understandable.	There are many grammatical errors. The expressions are difficult to understand.
Speed	The speed is very good.	The speed is good.	The speed is a little fast or slow.	The speed is very fast or slow.
Content	The student makes at least 8 sentences and the topic is attractive.	The student makes 8 sentences and the topic is interesting.	The student makes 8 sentences, but the topic is not so interesting.	The student makes 7 sentences or less even if the topic is interesting.
Presentation Skills	The tone of voice changes in a natural way.	The tone of voice changes.	The tone of voice changes a little.	The tone of voice sounds like a robot.
	The student's gestures and body language are lively and natural.	The student uses body language and gestures.	The student uses gestures and body language a few times.	The student barely moves. They are like a statue.
	We can hear easily.	We can hear.	We can't hear sometimes.	We can almost never hear.
	The student always looks at the audience.	The student mostly looks at the audience.	The student sometimes looks at the audience.	The student never looks at the audience.
Memorization	The student doesn't need any help and doesn't pause to remember.	The student doesn't need any help, but pauses to remember once or twice.	The student doesn't need any help, but pauses to remember several times.	The student needs help.

(3) 表 3 Interview Test は、授業中に ALT が実施している「インタビューテスト」の評価表である。生徒のプレゼンテーションテストの練習中に、予め定めた順番表に従い、実施している。

表 3 Group Interview Test: Student: Date:

	Excellent 10	Good 8	So-so 6	Poor 4
Grammar	Sentences are worded well and easily understandable.	Sentences are not grammatically correct, but understandable.	Sentences are not grammatically correct and a little hard to understand.	Student only uses one or two words at a time.
Vocabulary and expressions	The student understands the topic and uses many vocabulary	The student understands the topic but uses simple words	The student understands the topic a little and uses simple words	The student doesn't understand the topic
Language	Minimal Japanese is used		The student can understand with some Japanese	The student can't understand unless Japanese is used
Conversation	The student answers questions and asks interesting and relevant questions.	The student answers questions and asks questions.	The student answers questions but does not ask questions	The student cannot carry a conversation.
Time	Student makes good use of the allotted time		Student uses only about half the allotted time.	Student spends more of the allotted time in silence
Greeting	The student easily greets each other and has fun	The student greets each other	The student forgets greetings.	The student does not greet and is not having fun.
Eye contact	Eye contact happens often and feels natural.	The student makes eye contact.	The student makes eye contact sometimes.	The student does not make eye contact.
Confidence	The student is confident and seems relaxed.	The student is a little nervous	The student is shy and nervous but can answer questions.	The student is so shy and nervous they can't answer questions.
Asking for help	The student easily asks for help if they need it.	The student asks for help if they need it.	The student asks for help if they need it, but it is difficult for them.	The student does not ask for help if they need it.

4. テストの測定結果と考察

4.1 3つのテストの測定方法

筆者の勤務校では、縦列に評価項目を示し、横列にその評価基準と得点を示すルーブリックを使用している。また、生徒に教師と同じルーブリックを配布し、自己評価させている。生徒が自己評価得点を簡単に算出することができるように配慮して、教師も生徒も、ルーブリックの当てはまる評価基準の記述を○で囲み、横列に示した得点群（Excellent 10、Good 8、So-so 6、Poor 4）からそれぞれ当てはまる数値を選び、それらを合算し、その合計点を○の数で割り算して得られた数値の小数点第2位を四捨五入した数値を得点としてルーブリック評価表に記入する方式を採用した。テスト終了直後に教師は生徒から自己評価用ルーブリックを受け取り、教師と生徒間の得点差（教師得点－生徒自己評価得点＝得点差）を分析し、生徒のルーブリック評価基準を把握している。

本実践研究では、「英語会話」のクラスで行った3つのテスト（プレゼンテーションテスト、スピーチテスト、インタビューテスト）をこの方法で教師と生徒がそれぞれ評価して得られた結果をもとに、ルーブリック評価の成果と問題点について考察し、分析した。なお、教員の評価として以下の分析に用いた数値は、プレゼンテーションテストとスピーチテストについては日本人教員である筆者の評価を用い、インタビューテストは、ALTの評価を用いた。

英語能力の測定に用いた3つのテストの得点は、次のように収集した。

- (1) プレゼンテーションテスト： レッスン1（5月）、レッスン2（6月）、及びレッスン3（7月）でのルーブリック評価の得点を分析した。
- (2) スピーチテスト： 1回目（5月）、2回目（6月）、3回目（7月）でのルーブリック評価の得点を分析した。各クラスともひとつの授業時間内に4から5名ずつ実施したため、各回とも複数の授業を使用して実施した。
- (3) インタビューテスト： 1回目（5月）、2回目（6月）、3回目（7月）でのルーブリック評価の得点を分析した。

4.2 3つのテストの測定結果の考察

アウトプットテスト評価方法としてのルーブリック使用の成果を測る手段として一つは教師の得点を、そしてもう一つは生徒の自己評価得点を用いた。また、それらの二つの得点差を用いた。データの分析には、筆者が所有するマイクロソフト社製エクセル2016を用いた。

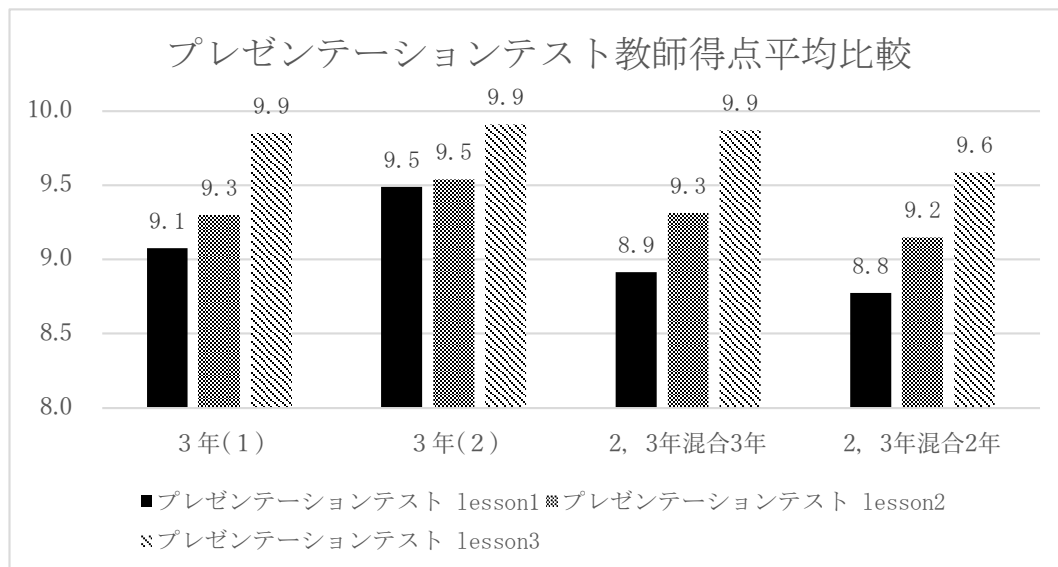
4.2.1 3つのテストにおけるルーブリック教師得点クラス平均

4.2.1.1 プレゼンテーションテストの教師得点平均（図表4）

4つの生徒群のそれぞれにおいて、レッスン1、レッスン2、レッスン3の3つの得点間で有意な得点の変化があったかどうかを見るために、一元配置の分散分析を行った結果、すべての生徒群でこれら3つの得点間に有意差が検出された（3年(1)： $F(2, 21)=14.877$, $p<.01$ 、3年(2)： $F(2, 27)=5.108$, $p<.05$ ）、2・3年混合3年： $F(2, 18)=8.152$, $p<.01$ 、2・3年混合2年： $F(2, 21)=13.332$, $p<.01$ ）。また、これらの生徒群のレッスン1と3の2つの得点間で両側t検定を行った結果、全ての生徒

群で有意差が検出された（3年(1)： $t(8)=-8.597$, $p<.01$ 、3年(2)： $t(10)=-3.771$, $p<.01$ 、2・3年混合3年： $t(7)=-5.404$, $p<.01$ 、2・3年混合2年： $t(8)=-10.942$, $p<.01$ ）。つまり、全ての生徒群で、レッスン1からレッスン3にかけての3ヶ月の間にプレゼンテーション能力の伸長があったことをルーブリックを使用した評価で見て取ることができた。

図表 4



プレゼンテーションテストクラス別平均点（最高10点 最低4点）

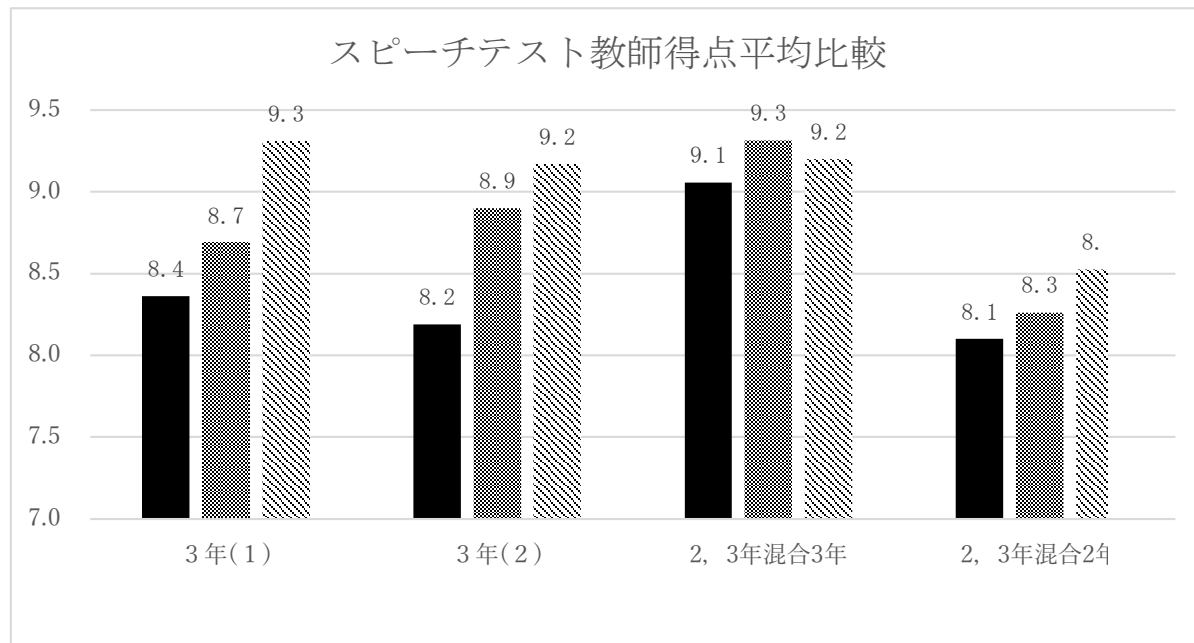
群	レッスン [実施日]	レッスン1	レッスン2	レッスン3
			3(1)：H28.5.9 3(2)：H28.5.9 2,3：H28.5.6	3(1)：H28.6.13 3(2)：H28.6.15 2,3：H28.6.17
3年(1)		9.1	9.3	9.9
3年(2)		9.5	9.5	9.9
2・3年混合3年		8.9	9.3	9.9
2・3年混合2年		8.8	9.2	9.6

4.2.1.2 スピーチテストの教師得点平均（図表 5）

4つの生徒群のそれぞれにおいて、1回目、2回目、3回目の3つの得点間で有意な得点の変化があったかどうかを見るために、一元配置の分散分析を行った結果、3年(1)と3年(2)において、これら3つの得点間に有意差が検出された（3年(1)： $F(2,21)=4.014$, $p<.05$ 、3年(2)： $F(2,27)=4.931$, $p<.05$ ）。また、これら2つの生徒群の1回目と3回目の2つの得点間で両側t検定を行った結果、いずれの生徒群においても有意差が検出された（3年(1)： $t(8)=-3.844$, $p<.01$ 、3年(2)： $t(10)=-6.515$, $p<.01$ ）。つまり、ルーブリックを使用した評価で見る限りスピーチテストに関しては、3年生の2群では3ヶ月間という短い期間でもスピーチ能力の

伸長を見て取ることができたが、2・3年混合群の3年生では、2年生とともに、スピーチ能力の伸長を見て取ることができなかった。

図表 5



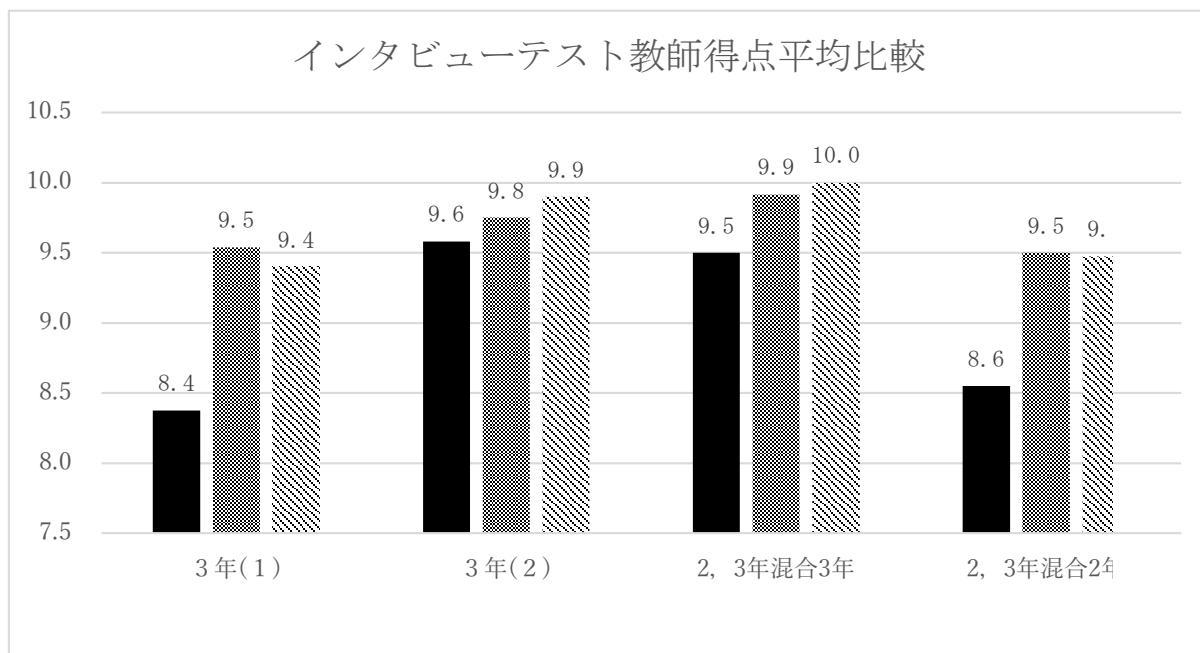
スピーチテストクラス別平均点 (最高 10 点 最低 4 点)

回 [実施日]	群		
	1回目 (3(1):H28.5.2~5.6 3(2):H28.5.9~5.11 2,3:H28.5.6~5.11)	2回目 (3(1):H28.6.3~6.6 3(2):H28.6.1~6.6 2,3:H28.6.1~6.15)	3回目 (3(1):H28.7.11~7.15 3(2):H28.7.4~7.11 2,3:H28.7.1~7.13)
3年(1)	8.4	8.7	9.3
3年(2)	8.2	8.9	9.2
2・3年混3年	9.1	9.3	9.2
2・3年混2年	8.1	8.3	8.5

4.2.1.3 インタビューテストの教師得点平均 (図表 6)

4つの生徒群のそれぞれにおいて、1回目、2回目、3回目の3つの得点間で有意な得点の変化があったかどうかを見るために、一元配置の分散分析を行った結果、3年(1)と2・3年混合2年で、3つの得点間に有意差が検出された(3年(1): $F(2, 21)=6.422$, $p<.01$ 、2・3年混合2年: $F(2, 21)=3.518$, $p<.05$)。しかし、3年(2)と2・3年混合3年では有意差が検出されなかった。3年(1)と2・3年混合2年の両クラスで1回目と3回目間の伸びが大きく、これら2クラスの1回目と3回目の2つの得点間で両側t検定を行った結果、両クラスとも有意差が検出された(3年(1): $t(8)=-2.440$, $p<.05$ 、2・3年混合2年: $t(8)=-2.691$, $p<.05$)。ルーブリックを使用した評価で見ると、これら2クラスの生徒群でインタビュー能力の伸長を見て取ることができた。

図表 6



インタビューテストクラス別平均点 (最高 10 点 最低 4 点)

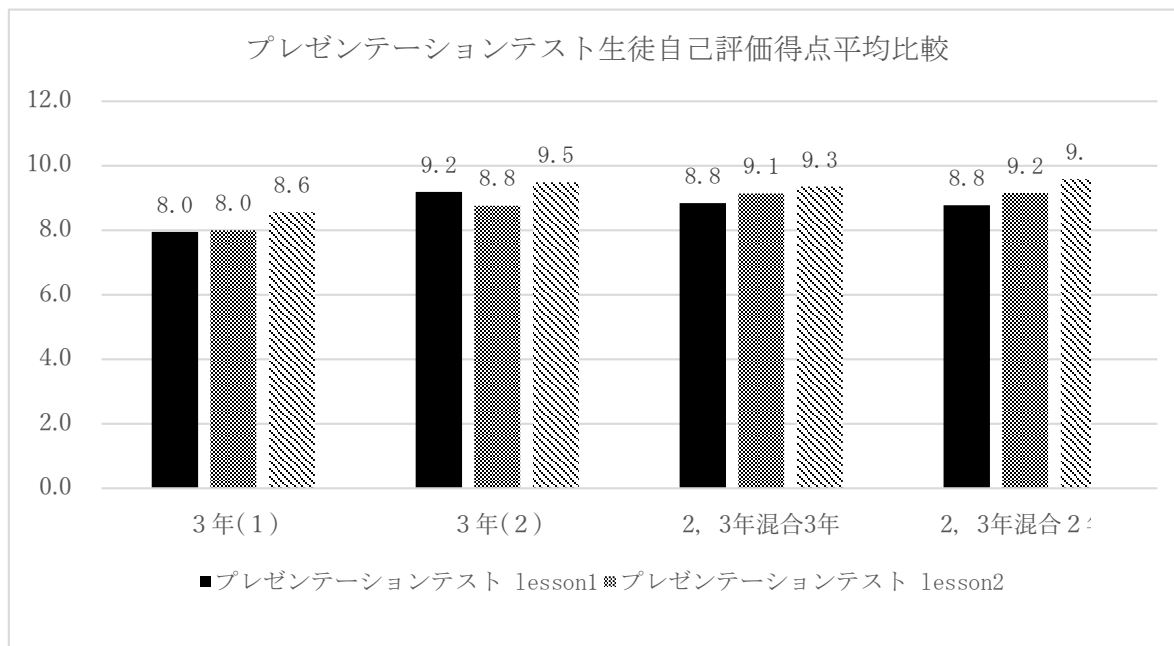
群	回 [実施日]	1回目 (3(1): H28. 5. 16 3(2): H28. 5. 16 2, 3: H28. 5. 20)	2回目 (3(1): H28. 6. 17 3(2): H28. 6. 20 2, 3: H28. 6. 22)	3回目 (3(1): H28. 7. 22 3(2): H28. 7. 20 2, 3: H28. 7. 20)
	3年(1)		8.4	9.5
3年(2)		9.6	9.8	9.9
2・3年混3年		9.5	9.9	10.0
2・3年混2年		8.6	9.5	9.5

4.2.2 3つのテストにおけるルーブリック生徒自己評価得点クラス平均

4.2.2.1 プレゼンテーションテストの生徒自己評価得点平均 (図表 7)

4つの生徒群のそれぞれにおいて、レッスン1、レッスン2、レッスン3の3つの得点間で有意な得点の変化があったかどうかを見るために、一元配置の分散分析を行った結果、どの群においても3つの得点間に有意差は検出されなかった。しかし、これら4つの生徒群のそれぞれにおいて、レッスン1と3の2つの得点間で両側t検定を行った結果、2・3年混合2年の生徒群で有意差が検出された ($t(8)=-2.627$, $p<.05$)。この生徒群の自己評価データは、5月から7月までの約3ヶ月間という短い期間で、緩やかに、しかし、確実にプレゼンテーション能力が伸長したことを示している。

図表 7



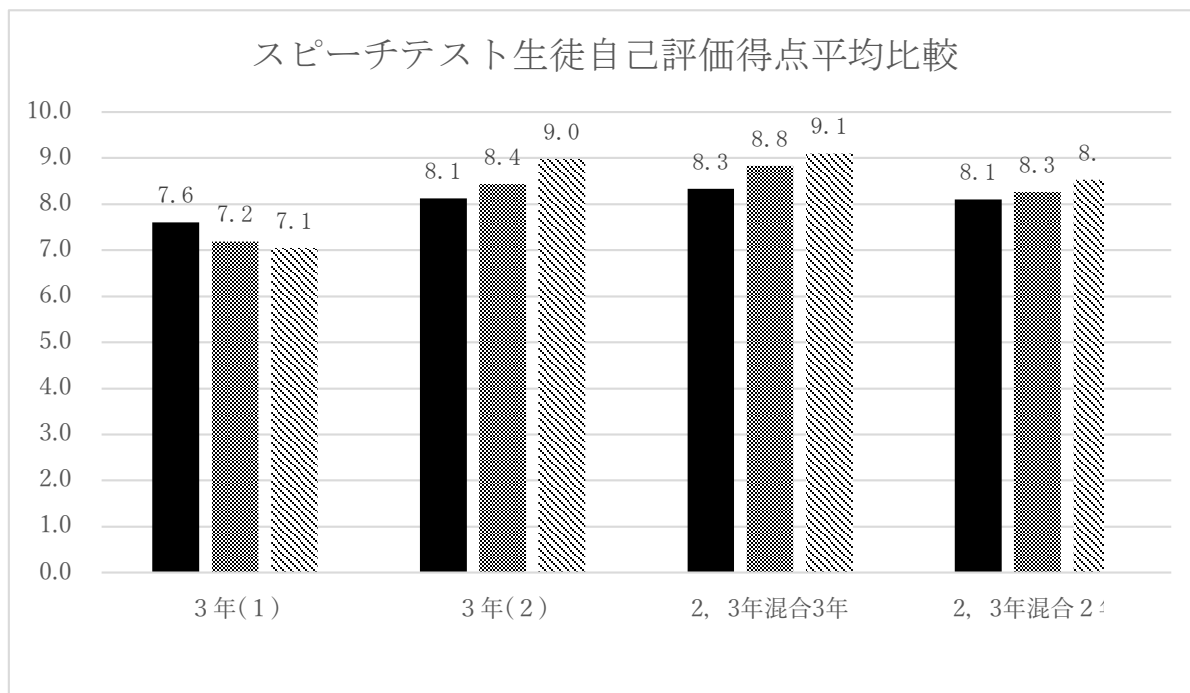
プレゼンテーションテストクラス別平均点 (最高 10 点 最低 4 点)

群	レッスン [実施日]	レッスン 1	レッスン 2	レッスン 3
			3(1): H28. 5.9 3(2): H28. 5.9 2, 3: H28. 5.6	3(1): H28. 6. 13 3(2): H28. 6. 15 2, 3: H28. 6. 17
3年(1)		8.0	8.0	8.6
3年(2)		9.2	8.8	9.5
2・3年混合3年		8.8	9.1	9.3
2・3年混合2年		8.8	9.2	9.6

4. 2. 2. 2 スピーチテストの生徒自己評価得点平均 (図表 8)

4つの生徒群のそれぞれにおいて、1回目、2回目、3回目の3つの得点間で有意な得点の変化があったかどうかを見るために、一元配置の分散分析を行った結果、それら3つの得点間に有意差は検出されなかった。しかし、これら4つの生徒群のそれぞれにおいて、1回目と3回目の2つの得点間で両側 t 検定を行った結果、3年(2)の生徒群で有意差が検出された ($t(10)=-2.524, p<.05$)。この生徒群の自己評価データは、3ヶ月間という短い期間で、緩やかに、しかし、確実にスピーチ能力が伸長したことを示している。

図表 8



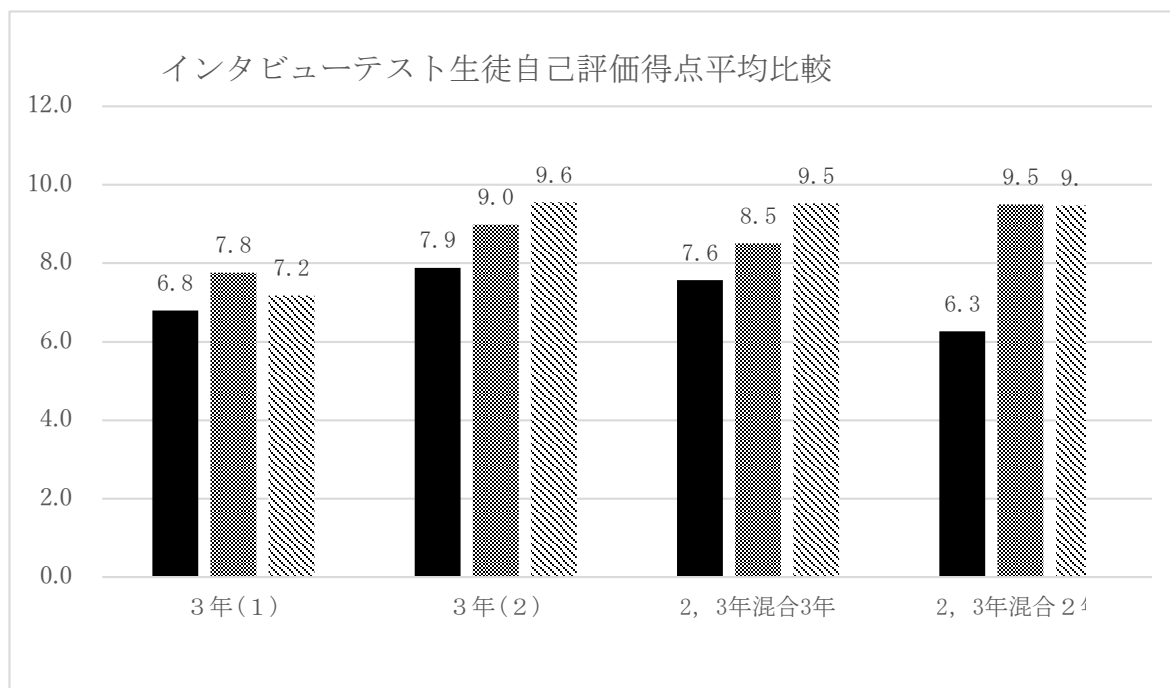
スピーチテストクラス別平均点 (最高 10 点 最低 4 点)

回 [実施日]	群		
	1回目	2回目	3回目
	3(1):H28. 5. 2~5. 6 3(2):H28. 5. 9~5. 11 2, 3: H28. 5. 6~5. 11	3(1):H28. 6. 3~6. 6 3(2):H28. 6. 1~6. 6 2, 3: H28. 6. 1~6. 15	3(1):H28. 7. 11~7. 15 3(2):H28. 7. 4~7. 11 2, 3: H28. 7. 1~7. 13
3年(1)	7.6	7.2	7.1
3年(2)	8.1	8.4	9.0
2・3年混3年	8.3	8.8	9.1
2・3年混2年	8.1	8.3	8.5

4. 2. 2. 3 インタビューテストの生徒自己評価得点平均 (図表 9)

4つの生徒群のそれぞれにおいて、1回目、2回目、3回目の3つの得点間で有意な得点の変化があったかどうかを見るために、一元配置の分散分析を行った結果、2・3年混合2年と3年(2)では有意差が検出された(2・3年混合2年: $F(2, 21)=4.200$, $p<.05$, 3年(2): $F(2, 17)=4.872$, $p<.05$)。また、これら2つの生徒群の1回目と3回目の2つの得点間で両側t検定を行った結果、3年(2)と2・3年混合2年の生徒群で有意差が、また、2・3年混合3年で有意傾向が検出された(3年(2): $t(10)=-3.929$, $p<.01$, 2・3年混合3年: $t(7)=-2.212$, $0.05<p=0.06<0.1$, 2・3年混合2年: $t(8)=-3.950$, $p<.01$)。これらの生徒群の自己評価データは、5月から7月までの3ヶ月間という短い期間で、緩やかに、しかし、確実にインタビュー能力が伸長したことを示している。

図表 9



インタビューテストクラス別平均点 (最高 10 点 最低 4 点)

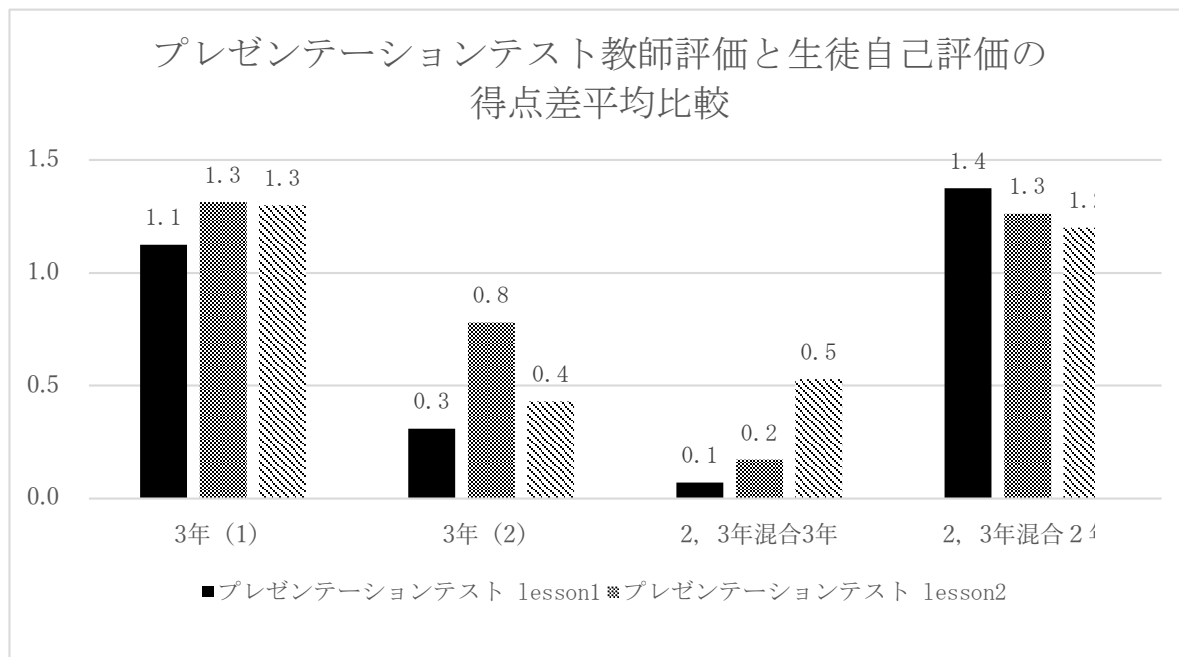
回 [実施日]	群		
	1回目 (3(1): H28. 5. 16 3(2): H28. 5. 16 2, 3: H28. 5. 20)	2回目 (3(1): H28. 6. 17 3(2): H28. 6. 20 2, 3: H28. 6. 22)	3回目 (3(1): H28. 7. 22 3(2): H28. 7. 20 2, 3: H28. 7. 20)
3年(1)	6.8	7.8	7.2
3年(2)	7.9	9.0	9.6
2・3年混3年	7.6	8.5	9.5
2・3年混2年	6.3	9.5	9.5

4.2.3 3つのテストにおける教師と生徒間のルーブリック得点差平均

4.2.3.1 プレゼンテーションテスト得点差平均 (図表 10)

レッスン 1、レッスン 2、レッスン 3 の 3 つのテストを行うにつれて、生徒がルーブリックによる評価方法に慣れ、教員と同じ基準で客観的に自分のパフォーマンスを評価できるようになったのだろうか。教員評価の得点と生徒の自己評価の得点の間の差の変化を見て、その差が減少傾向を示せば、そのような変化が起こったかどうかを知ることができる。そこで、4 つの生徒群のそれぞれにおいて一元配置の分散分析を行った結果、どの生徒群でも、3 つのテストにおける教員評価と生徒の自己評価の間の差を示す 3 つ値の間に有意差は検出されなかった。4 群の中で、唯一 2・3 年混合 2 年の生徒群だけが、レッスン 1 での差 (1.4) からレッスン 3 での差 (1.2) にかけて減少しているため、その 2 つの得点差間で両側 t 検定を行ってみたが、結果として有意差は検出されなかった。3 ヶ月間という短い期間ではその差が減少したことを示す有意なデータを収集することができなかった。

図表 10



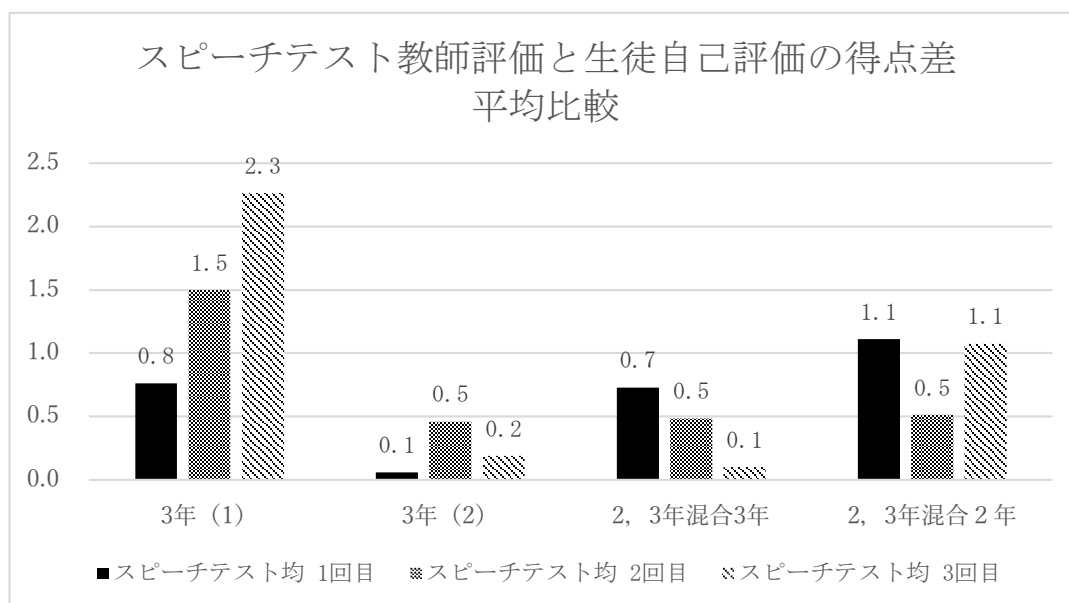
プレゼンテーションテストクラス別得点差平均

群	回	1回目	2回目	3回目
	3年(1)		1.1	1.3
3年(2)		0.3	0.8	0.4
2・3年混合3年		0.1	0.2	0.5
2・3年混合2年		1.4	1.3	1.2

4.2.3.2 スピーチテスト得点差平均 (図表 11)

前節で述べたのと全く同じ目的で、1回目、2回目、3回目のスピーチテストにおける教員評価と生徒の自己評価の得点差を示す3つの値の間で一元配置の分散分析を行った結果、4つの生徒群のいずれにおいても有意差は検出されなかった。ただ、4群の中で唯一2・3年混合3年の生徒群においてのみ1回目(0.7)と3回目(0.1)の間の差が減少しているため、その2回のテストの評価得点差間で両側t検定を行ってみた。その結果、有意差は検出されなかった。プレゼンテーションテスト同様、スピーチテストにおいても、3ヶ月間という短い期間ではその差が減少したことを示す有意なデータを収集できなかった。

図表 11



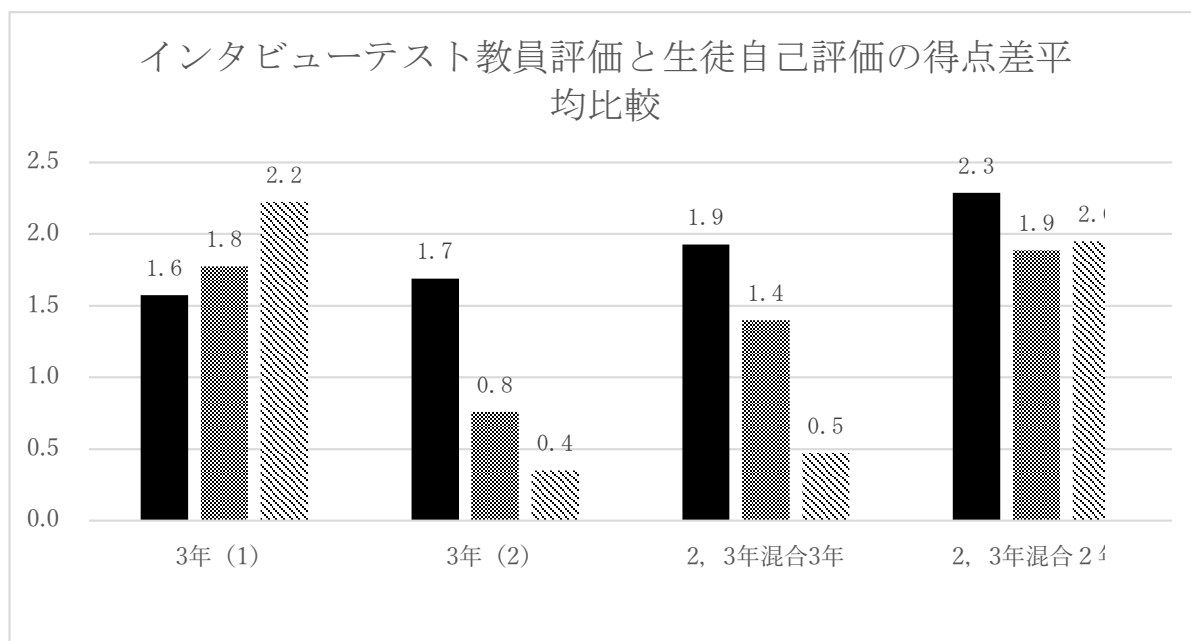
スピーチテストクラス別得点差平均

群 \ 回	1回目	2回目	3回目
3年(1)	0.8	1.5	2.3
3年(2)	0.1	0.5	0.2
2・3年混合3年	0.7	0.5	0.1
2・3年混合2年	1.1	0.5	1.1

4.2.3.3 インタビューテスト教師評価と生徒自己評価の得点差平均 (図表 12)

これもまた、前節で述べたのと全く同じ目的で、1回目、2回目、3回目のインタビューテストにおける教員評価と生徒の自己評価の得点差を示す3つの値の間で一元配置の分散分析を行った結果、3年(2)の得点差間に有意差が検出された ($F(2, 27)=3.985, p<.05$)。また、3年(2)と2・3年混合3年及び2年の生徒群では、1回目と3回目の間の得点差が減少しているため、これら3つの生徒群において、1回目と3回目の教員評価と生徒の自己評価の得点差間で両側t検定を行った結果、3年(2)では有意差が、2・3年混合3年では有意傾向が検出された (3年(2): $t(10)=3.603, p<.01$ 、2・3年混合3年: $t(7)=2.386, 0.05<p=0.054<0.1$)。しかし、2・3年混合2年では有意差が検出されなかった。したがって、インタビューテストでは、プレゼンテーションテストとスピーチテストとは異なり、3ヶ月という短い期間でその差が減少したことを示す生徒群が出た。

図表 12



インタビューテストクラス別得点差平均

群 \ 回	1回目	2回目	3回目
3年(1)	1.6	1.8	2.2
3年(2)	1.7	0.8	0.4
2・3年混合3年	1.9	1.4	0.5
2・3年混合2年	2.3	1.9	2.0

4.2.4 英語に対する自信の度合いに応じて2つの生徒群に分けた場合の3つのテストの測定結果

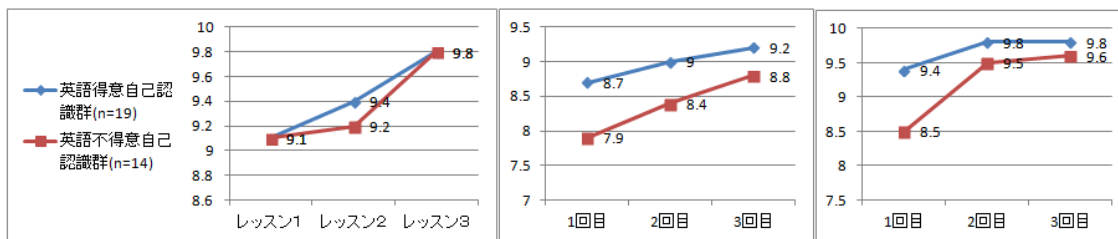
8月22日～8月26日にかけて授業の一部を使って、上で述べてきた4つの生徒群の全員に対して英語会話の授業に関するアンケートを実施した。そのアンケートの中で、「英語が得意ですか。」という質問項目に「得意である」、「まあまあ得意である」、「あまり得意でない」、「得意でない」という4つの選択肢から一つを選ぶ形式で回答させた。その結果、「英語が得意である」または「まあまあ得意である」と自己認識している「英語得意自己認識群」の構成は男子6名、女子13名、計19名、「あまり得意ではない」または「得意でない」と自己認識している「英語不得意自己認識群」の構成は男子9名、女子5名、計14名であった。

4.2.4.1 3つのテストにおける「英語得意自己認識群」と「英語不得意自己認識群」のルーブリック教師得点平均 (図表 13)

レッスン1、レッスン2、レッスン3(プレゼンテーションテスト)または1回目、2回目、3回目(スピーチテスト及びインタビューテスト)の3つの得点差間で一元配置の分散分析を行った結果、英語不得意自己認識群では3種すべてのテストで3値の間に有意差が検出されたが、英語得意自己認識群で有意差があったのはプレゼ

ンテーションテストだけで、スピーチテスト及びインタビューテストでは有意傾向にとどまった。

図表 13 3種アウトプットテストの教師得点平均

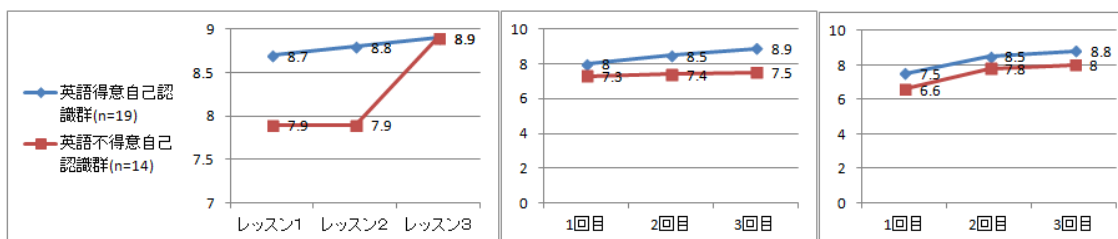


群	プレゼンテーションテスト			スピーチテスト			インタビューテスト		
	レッスン1	レッスン2	レッスン3	1回目	2回目	3回目	1回目	2回目	3回目
英語得意自己認識群 (n=19)	9.1	9.4	9.8	8.7	9.0	9.2	9.4	9.8	9.8
	F(2, 54)=17.915, p<.01								
英語不得意自己認識群 (n=14)	9.1	9.2	9.8	7.9	8.4	8.8	8.5	9.5	9.6
	F(2, 39)=4.058, p<.01			F(2, 39)=5.912, p<.01			F(2, 39)=9.212, p<.01		

4.2.4.2 3つのテストにおける「英語得意自己認識群」と「英語不得意自己認識群」のルーブリック生徒自己評価得点平均 (図表 14)

レッスン1、レッスン2、レッスン3 (プレゼンテーションテスト) または1回目、2回目、3回目 (スピーチテスト及びインタビューテスト) の3つの得点差間で一元配置の分散分析を行った結果、「英語得意自己認識群」ではインタビューテストで、「英語不得意自己認識群」ではプレゼンテーションテストとインタビューテストで ($F(2, 39)=4.058, p<.05$; $F(2, 39)=9.217, p<0.01$)、それぞれ有意差が検出された。

図表 14 3種アウトプットテストの生徒自己評価得点平均



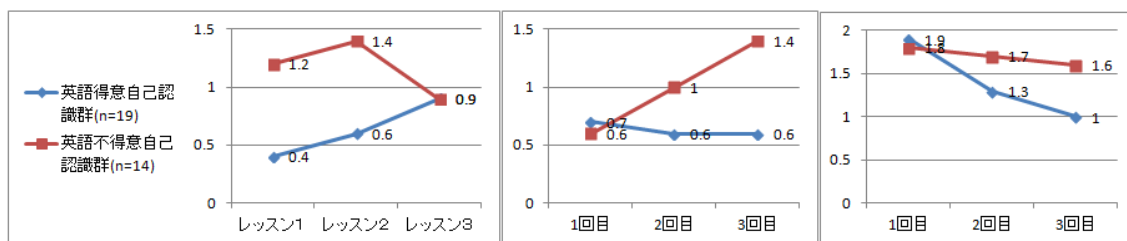
群	プレゼンテーションテスト			スピーチテスト			インタビューテスト		
	レッスン1	レッスン2	レッスン3	1回目	2回目	3回目	1回目	2回目	3回目
英語得意自己認識群 (n=19)	8.7	8.8	8.9	8.0	8.5	8.9	7.5	8.5	8.8
							F(2, 54)=4.584, p<.05		
英語不得意自己認識群 (n=14)	7.9	7.9	8.9	7.3	7.4	7.5	6.6	7.8	8.0
	F(2, 39)=4.058, p<.05						F(2, 39)=9.217, p<.01		

4.2.4.3 3つのテストにおける「英語得意自己認識群」と「英語不得意自己認識群」の教師と生徒間のルーブリック得点差平均 (図表 15)

レッスン1、レッスン2、レッスン3 (プレゼンテーションテスト) または1回目、

2回目、3回目（スピーチテスト及びインタビューテスト）の3つの得点差間で一元配置の分散分析を行った結果、有意差が検出されたのは「英語得意自己認識群」のインタビューテストだけだった($F(2, 54)=3.416, p<.05$)。また、このテストの初回と最終回の2つの得点間で両側t検定を行った結果、有意差が検出された($t(19)=2.382, p<.05$)。インタビューテストでは、「不得意群」には有意差がなかったものの、どちらの生徒群も教師との評価得点差が縮まる傾向を示した。ところが、プレゼンテーションテストとスピーチテストではこの2群が反対の傾向を示した。プレゼンテーションテストでは「得意群」が、教師評価との差を縮める傾向を、「不得意群」は差が広がる傾向を示した。他方、スピーチテストでは、「得意群」はわずかに減少傾向を示したが、「不得意群」は差が大きく広がった。英語を得意と感じているかどうかという意識の差と3種のアウトプットテストの差という二つの要因が一体どう影響しあっているのかわからないが、今後の研究課題としたい。

図表 15 3種アウトプットテストの教師評価と生徒自己評価の得点差平均



群	レッスン/回	プレゼンテーションテスト			スピーチテスト			インタビューテスト		
		レッスン1	レッスン2	レッスン3	1回目	2回目	3回目	1回目	2回目	3回目
英語得意自己認識群 (n=19)		0.4	0.6	0.9	0.7	0.6	0.6	1.9	1.3	1.0
英語不得意自己認識群 (n=14)		1.2	1.4	0.9	0.6	1.0	1.4	1.8	1.7	1.6

4.3 成果と問題点

本実践研究では、限定的ではあるが、「英語会話」授業におけるアウトプット評価方法としてのルーブリック使用の成果を可能性として示すことができたのではないかと考える。特に、3ヶ月間という短い期間で生徒のプレゼンテーション能力を順調に伸ばすことができた。この能力伸長をもたらした要因が何かということ、客観的なデータの形では明示できないが、教室現場の観察情報を加えることが許されるなら、ルーブリック評価項目に tone (口調)、expression (表情)、そして gesture (ジェスチャー) を設定したことが功を奏した結果であると断言してよい。生徒は臨場感あふれるパフォーマンスを初回から見せてくれていた。また、同様に、スピーチテストとインタビューテストでもルーブリックにコミュニケーションスキルに関する評価項目を設定したことが功を奏したと判断する。テスト中、私の注意を待つまでもなく、ルーブリックを意識して、目線をきちんと聴衆または話し相手に向けて表情豊かに話す生徒の姿を見ることができた。また、「英語不得意自己認識群」の生徒達が頑張っ、3ヶ月間という短い期間で、3つのテストでスピーキング能

力の大きな伸長を見てとることができたのも、このルーブリックによる自己評価によるところが大きいと考える。

しかし、この伸長がルーブリックを使った自己評価を授業活動として導入したことだけに帰することができるかと言えば、もちろん否である。自己評価活動をさせない統制群またはルーブリック評価以外の自己評価活動をさせる対照群を使って、今回見たような3種のアウトプットにおける技能の伸長を比較しなければならない。

また、別の課題も残った。教師得点と生徒自己評価得点間の差が回を重ねるごとに縮小したことを示す有意なデータを収集しなかったが、4つの生徒群のうち1つ、3年(2)のインタビューテストにおける評価で両者の間の差が有意に縮小したことが示されたにとどまった。両者の差が縮まらなかった要因のひとつは、生徒の自己評価が予想以上に厳しかったことである。教師評価と生徒自己評価のすり合わせを授業中に時間を設けて生徒にさせなかったことが主な要因なのか、それとも他の要因が影響したのかについて今後検証しなければならない。さらに、なぜインタビューテストだけ、また、なぜ3年(2)の生徒群だけ、生徒の自己評価と教員評価の間の差が有意に縮まったのかについても今後の課題としたい。

5. おわりに

教師と学習者が同じルーブリックを使用して、双方が学習者のアウトプットの評価をすることは、教師から生徒へという評価の一方通行的な流れを双方向の流れに変えることに繋がる。つまり、教師は学習者の自己評価を見て学習者のルーブリック評価基準の理解度を確認する機会が与えられるだけでなく、自分の評価の正確性について見直す機会も与えられる。なぜならば、教師はアウトプットテストにおける学習者の実技を見て即座に評価しなければならないため、各テストで同じ基準で評価しなければならないにも関わらず、時には厳しく評価したり、あまく評価したりする可能性があるからである。また、学習者は教師の評価を見て自己の良い点と改善点を確認する機会を与えられるだけでなく、自己評価の正確性について見直す機会も与えられる。このようにルーブリックの使用は教師と学習者の双方にとって有益なことであるので、今後さらにルーブリックを使った評価についての研究を進めていきたいと考える。

引用文献

- Young, F. & Wilson, R. J. (2000). *Assessment & learning: ICE approach*. (土持ゲアリー法一 (監訳)、小野恵子 (訳) (2013) 『「主体的学び」につなげる評価と学習方法—カナダで実践される ICE モデル—』東信堂)
- 鈴木孝明・白畑知彦 (2013) 『ことばの習得—母語獲得と第二言語習得』くろしお出版
- 當作靖彦・中野佳代子 (2013) 『外国語学習のめやす—高等学校中国語と韓国語教育からの提言』公益財団法人 国際文化フォーラム

はじめてルーブリック評価を経験した日本人大学生英語学習者の評価及びフィードバックに見るルーブリック評価実施上の留意点(2)

渋谷 奈緒美
(北海道平取高等学校)

Some Issues in Using Rubrics in the Classroom: Assessment and Survey Results with Japanese College Learners of English with No Previous Experience with Rubric Assessment, Part 2

Naomi SHIBUYA
(Sapporo Otani University, Part-time Faculty)

Abstract

This paper is the second part of a report on the survey results concerning classroom use of rubrics as an assessment tool. The first report discussed 22 out of 25 college freshman students taking required general English courses who belonged to higher proficiency brackets in a group of 50 students that was surveyed in the 2016 academic year. This time, test data as well as questionnaire responses were collected from 23 students out of the other 25 who were relatively lower in proficiency than the first group of 25 students. At the end of a 15-week semester, the 23 freshman students were given a performance test using rubrics and responded to a survey on the test. Both the test data and survey results, combined with those from the previous survey, revealed and shed light to some issues pertinent to those who plan to implement rubric assessment in the college classroom. They suggested (1) one-time, instead of frequent, use of self and peer reviews is not good enough to both raise student confidence in rubric assessment as well as their ability to evaluate their own performance and motivate them to attain higher goals; (2) teachers should regularly and gradually help students calibrate their self-assessment so it might approximate teacher assessment; and that (3) the assessment method, such as videotaping, that would help both students and teachers assess the recorded performance as objectively as possible might give students, especially those in lower proficiency brackets, undue stress, which works negatively against successful rubric assessment.

1 はじめに

この論文は、大学の英語授業でのルーブリックの活用が広まる一助となることを目的として、学生のパフォーマンスをルーブリックで評価する際の問題点をあぶり出したパート1の続編である。パート1では、非言語系の学科で英語を学ぶ大学1年生の集団を研究対象として、入学時に行われた習熟度測定テストの結果によって4クラスに分けられたうちの上位2クラスの学生にルーブリックを用いたインタビュー形式の実技試験を課し、試験後のアンケート結果と、試験1週間前に実施した模擬試験での学生の自己評価及び実際の試験での教員評価の結果を分析した。パート2では、パート1と同じ学科で学ぶ大学生で習熟度が下位2クラスの学生を比較研究対象として、ルーブリック評価を行う際に注意すべき点をさらにあぶり出し、パート1での研究を発展させた。

パート1では、1度限りのルーブリック評価では、ルーブリック評価の持つ利点を生かしきれないということを報告した。本番となる試験前に模擬試験を行い、その様子を動画として撮影させ、それを見ながら自分のパフォーマンスを客観的に判断する練習をしても、学生の自己評価は教師の評価と同じレベルには達しなかった。また、圧倒的多数の学生が会話能力の測定方法としてルーブリック評価を肯定的に受け止め、その有益性を認める一方で、多くの学生がビデオ撮影に抵抗感を覚え、ビデオ撮影されると、普通に話す場合よりも緊張して、実力が発揮できない気がする」と答えており、1度限りの評価の方法としてビデオ撮影を用いることが試験でのパフォーマンスに否定的な効果を生じさせる可能性があることを示唆した。

そこで、このパート2では、習熟度の低い下位2層を検証し、習熟度の高い上位2層とルーブリック評価を実施する際に配慮すべき点は同じなのか、それとも異なるのか、また、自分の英語能力を正しく把握して、自ら適切な目標設定とそれに向かう努力を自主的にすることができるようになるというルーブリックの効果を引き出すためには、やはり1度限りの評価練習では無理なのかなどについて調べることにした。パート1に続いて、筆者が非常勤講師を務めていた札幌大谷大学の理解と協力を得て、実際の科目指導の中で引き続き実践研究をさせていただいた。以下その研究の結果を報告する。

2 方法

2.1 実施クラスと試験までの手順

パート1では、非言語系学科の大学1年生の必修科目を4つの習熟度別クラスで履修している学生のうち最上位クラス（以下「a」又は「aクラス」）10名（“All in English”クラス）と次位クラス（以下「b」又は「bクラス」）12名の計22名を研究の対象とした。この2クラス（aクラスとbクラス）でルーブリックを用いた試験を実施し、試験終了後に、ルーブリック評価及びビデオ撮影を用いた学習に関するアンケートを授業内で行った。

一方、今回の研究対象は、4つの習熟度別クラスの上から3番めのクラス（以下「c」又は「cクラス」）12名と4番目（最下層）（以下「d」又は「dクラス」）11

名の計 23 名とした。実際の科目履修者数は a クラス 12 名、b クラス 14 名の計 26 名だったが、アンケートデータが欠損していた 4 名をアンケートの分析から除外した。同様に、c クラスと d クラスの実際の履修者数はそれぞれ 12 名ずつの計 24 名だったが、パート 1 と同じ理由で、アンケートに回答しなかった d クラスの 1 名をアンケートの分析から除外した。そして、パート 1 で上位 2 層 22 名に行ったのと同じアンケートを下位 2 層の 23 名に実施し、下位 2 層 (c と d) 間の比較と、上位 2 層 (a、b) と下位 2 層 (c、d) 間の比較と、全 4 層間の比較を行った。

また、ルーブリックを用いた本試験で教員が各学生に与えた評価と、その 1 週間前に同じルーブリックを用いて行った模擬試験で学生が自分自身につけた自己評価との間で、アンケートと同様の各種比較を行った。評価値の比較においては、模擬試験と本試験のいずれか一方又は両方を受験しなかった学生と自己評価表を提出しなかった学生を分析から除外したため、全 4 層 50 名のうち、a クラス 9 名、b クラス 8 名、c クラス 10 名、d クラス 10 名の計 37 名のみを分析対象とした。

2.1.1 ルーブリックを用いた試験の概要と本研究の主なねらい

試験内容は、後期学習項目のうち、「レストランで注文をとる」について、ウェーター・ウェイトレスとして客の注文をとる試験を実施した。半期 15 回の授業のうちの 14 回目に、試験にはルーブリック評価表を使い評価することを告げ、評価表を配布した。パート 1 と同様に、第 14 回目の授業では、90 分の授業のうち、まず前半 45 分を、ルーブリック評価表を見ながらメニュー表を使い「ウェーター・ウェイトレスとして注文をとる」という課題に向けた練習をする時間に充てた。そして、残りの後半 45 分は、スマートフォンのムービーで、メニュー表を使いながら実際の試験を想定して注文をとっている場面を撮影し、(1)ルーブリックによる自己評価、(2)ペアの評価、(3)自身のパフォーマンスを撮影したムービーの 3 点を持ち帰り、復習して 1 週間後の試験に備えるように指示をした。また、15 回目にあたる試験日に、同じルーブリック評価表を用い、メニュー表を別のものに差し替えて試験を実施した。スマートフォンでの撮影は、パフォーマンスの評価を正確に行うためのパフォーマンスの記録および評価の確認方法として當作 (2014) が紹介した方法である。パート 1 と同じ環境のもとで試験をするために、今回も導入した。

當作・中野 (2013) が示しているルーブリック評価表の使い方によれば、評価表は期頭に学生に提示し、授業内で使い、何度かの評価練習で自己の内省を具体的な数値として表現する経験をさせることが前提となっている。そのような過程を経て本番の試験で使用することではじめて、指導と評価の両方の役割を同時に果たすことのできるルーブリック評価の効果が得られるものと考えられている。しかし、本研究のパート 1 で行った教育実践では、諸般の事情からそのような適切な手順を踏むことができず、学生がルーブリック表を用いた評価を行ったのはたった 1 度限りだった。その結果、1 度限りのルーブリック評価を経験するだけでは、所期の効果は得られないことが示唆された。後期のパート 2 では、習熟度の下位層を対象として全く同じアンケートを取り、上位層と下位層で何らかの差異が生じるのか、また、母体を厚くしたことにより分析結果に変化が生じるのかを調査する計画を立て、前回と同じ条件の下でルーブリック評価を行った。本研究では、上に述べた試験のデ

ータと試験後に実施したアンケートで得られたデータをもとに、次の3点について調査を行った。

- (1) 1度限りの練習の直後にインタビュー試験を経験させた場合、学生たちはルーブリック評価という評価方法をどのように受け止め、認識するのか。また、習熟度によってその受け止め方に差はあるのか。
- (2) パート1の結果では、習熟度が上位のクラスでは、ルーブリック評価の練習が1度限りだったために、学生の自己評価と教員の評価との間には有意な差が生じた。ルーブリック評価の練習が1度限りの場合、習熟度が下位のクラスでも同じ結果となるのか。
- (3) 習熟度が上位のクラスに、ルーブリックによる自己評価の練習として、本試験前に自身のパフォーマンスの様子をスマートフォンで撮影し、ルーブリック評価表を参照して自己評価をする課題を本試験の前に与えたが、その回数を十分に行わないことが学習者には負担やストレスの原因となり、逆効果を生む傾向が認められた。下位のクラスでもその傾向が認められるのか。

2.1.2 評価に使用したルーブリック

評価には、図表1、2、3のルーブリックを使用した。

図表1 「語彙・単語」のルーブリック

10	8	6	4	0
レストランで注文を取る表現を身につけている。	レストランで注文を取る表現を、ほぼ覚えているがあいまいなものもある。	レストランで注文を取る表現を半分ほどしか覚えていない。	レストランで注文を取る表現をほぼ覚えられていない。	レストランで注文を取る表現を覚えていない。

図表2 発音・アクセント・強弱・リエゾン

10	8	6	4	0
発音やアクセントを正しく発音できる。日本語のようなまっすぐな読み方ではなく、英語の強弱やリエゾンに注意している。	発音やアクセントに若干の問題があるが、日本語のまっすぐな読み方ではなく、英語の強弱やリエゾンに注意しようとしている。	発音にたびたび間違いが見られるが、日本語のまっすぐな読み方ではなく、英語の強弱やリエゾンに注意しようとしている。	発音に自信がないようである。読み方も日本語のようにまっすぐだが、なんとか強弱をつけようと努力している。	発音に自信がないようである。読み方は日本語のようで、強弱がなく、まっすぐ読みである。練習が必要である。

図表3 「態度」のルーブリック

30	25	20	10	0
レストランで注文を取る時、お客様が不快にならないスムーズな会話ができる。サラダからデザートまで5品目の注文は取ることができた。	注文を取る際の英語はほぼ適切であるが、話すことにせいっぱいで態度まで気を遣う余裕がない。会話がとぎれてしまうが、サラダからデザートまでの5品目の注文は取ることができた。	レストランで注文を取る英語に不安があり、なんとか頑張っ持っている表現を使いながら注文を取っている。かろうじて通じ、なんとか注文をとることができた。	レストランで注文を取る時の表現を覚えておらず、自分の持っている語彙でなんとか伝えようとはしているものの、注文がしっかりと取れないものもあった。	レストランで注文を取る時の表現を覚えておらず、注文を取ることができない。注文を取る時の英語を覚えなおす必要がある。

「語彙・単語」は0点から10点までを5段階で評価した。「発音、アクセント、強弱、リエゾン」も、平常の授業の様子から学生の差が大きいことから配点を大きくするのは適切ではないと考えて0点から10点までを5段階で評価した。「態度」など表現については、「レストランで注文をとる」場面の語彙を十分に覚えた上で、ウェイター・ウェイトレスとしてどのようなパフォーマンスをすればサラダからデザートまでの5品目の注文を取ることができるかを見るもので、「態度」という基準の意味するものが学生に十分理解されるように、テストの前の練習時に教師が例を示した。言語項目の正確さよりも課題を遂行することそのものに重点を置くことを学生に対して明確にするために、他の2つの基準の3倍の重みをつけ、計30点の配点にした。ただし、統計分析に当たっては、上の3つの評価領域の得点を同じ尺度に統一するため、「態度」の30-25-20-10-0という5段階評価は10-8-6-4-0とみなして分析した。

2.2 試験後のアンケート

アンケートでは、ルーブリックを使用した試験(質問1)、試験前のビデオ撮影(質問2)、事前に行った1時間の練習(質問3)、リスニングを含む筆記による試験(以下「筆記試験」とルーブリックを用いたインタビュー試験(質問4)、ルーブリックのような評価表(質問5)、ルーブリックとビデオ撮影を用いた試験(質問6)、事前学習時の評価をペアの相手と一緒にしたこと(質問7)、今後もルーブリックを用いた試験を受けたいと思うかどうか(質問8)の8項目について選択式で回答を求めた。また、質問9として、「最後に、今回のルーブリック評価によるインタビュー形式試験を受けてみて、良かった点や改善すべき点(こうしてもらえればもっと良かったと思うことなど)があったら、自由に書いてください。」という自由記述項目を設けた。選択式の質問では、質問1と質問3を除く全ての項目について4件法(とてもそう思う/少しそう思う/あまりそう思わない/全くそう思わない)で回答を求めた。

2.3 統計処理について

統計処理には IBM SPSS Statics Base V.21 を用いた。

3 結果及び考察

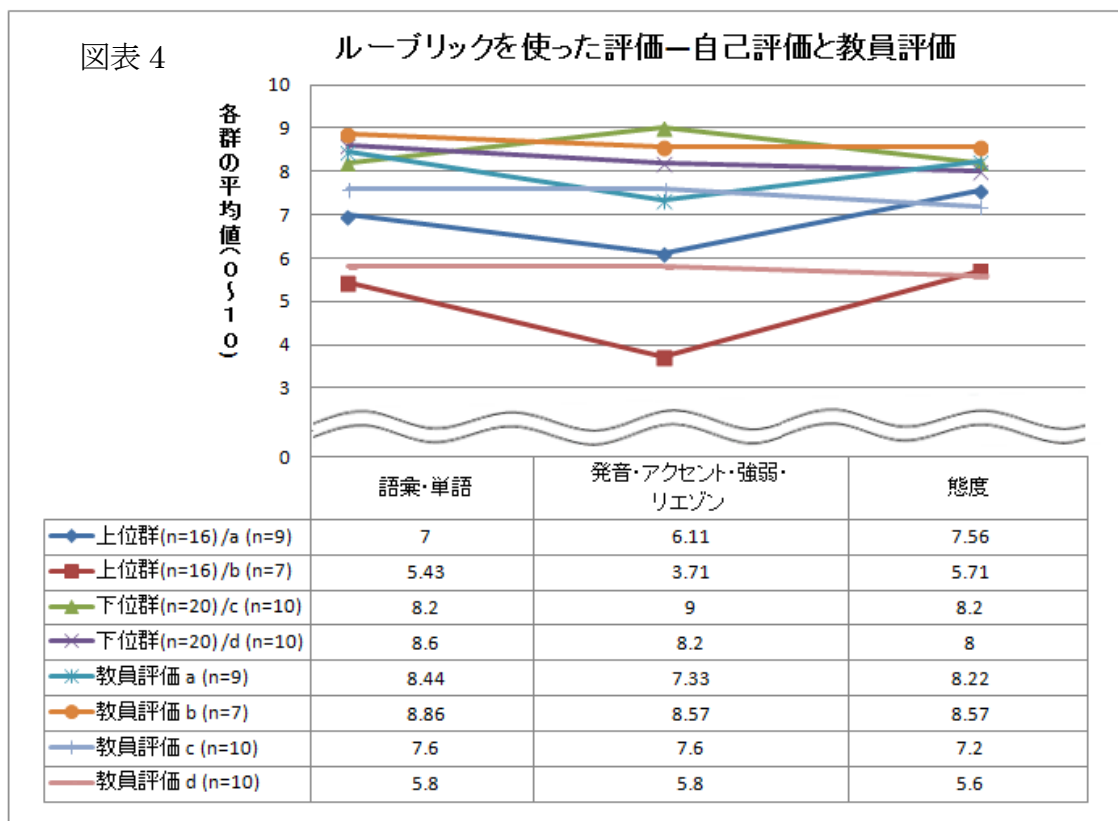
今回の実践研究は、下位2群(cクラス・dクラス)を直接の研究対象とするものであったが、この下位2群の間の差異だけではなく、上位2群と下位群の間、あるいは、全4群の間の差異が明らかになるように、以下では、今回の調査結果と前回の上位2群についての調査結果を併せて報告する。

3.1 ルーブリックを用いた試験の結果

3.1.1 自己評価と教員評価の一致度：学生はルーブリックにしたがって、自分の能力を適切に評価できたか

下位層c・d両クラスを含む全4クラスでの、予行練習時の学生自身の自己評価と、教師が1週間後の15回目に行ったパフォーマンス試験時の教師の評価は、図表4の通りとなった。1週間という実施時期の差を無視し、学生の自己評価と教師の

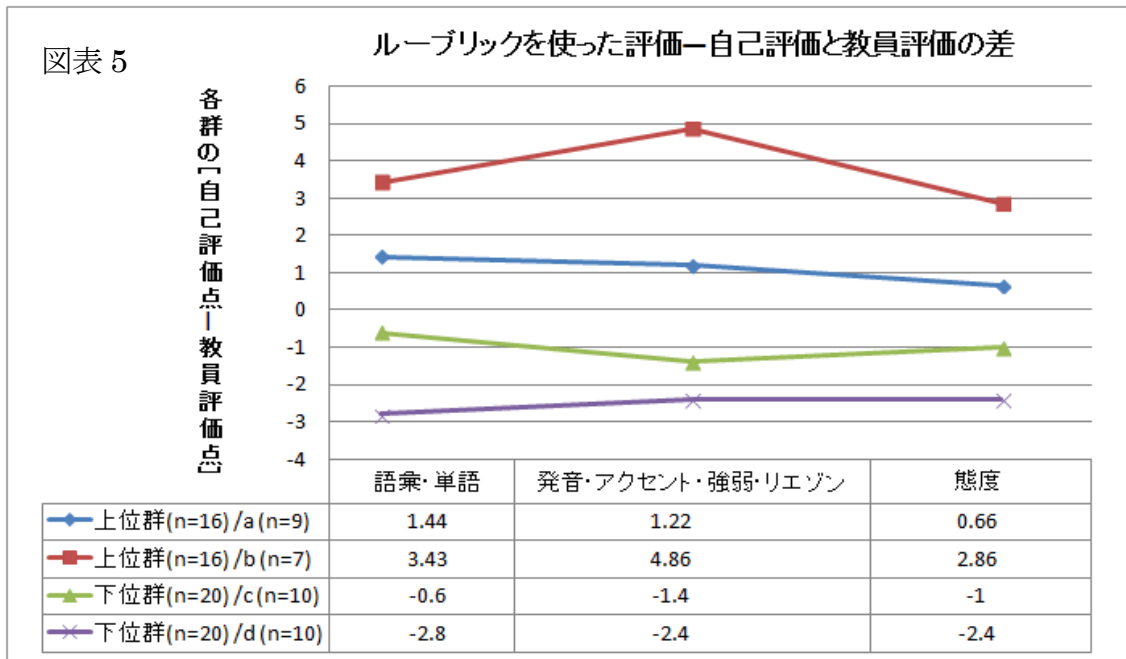
評価が同時期に行われたものとみなして両者の平均を t 検定にかけて比較してみると、3 つの評価領域のいずれにおいても有意差が検出され、同等とはみなせないことがわかった (a、b、c、d それぞれ $p < 0.05$)。ルーブリックを用いることで、教員と学生との間で評価基準についての認識が完全に共有されれば、結果として評価点数も教員と学生との間で同じような値を取るはずだが、実際はそうはならなかったということが示された。



次に、学生の自己評価と教員評価がどれだけ違っていたか、その差だけに注目してみると、図表 5 のようになった。

図表 5 の数値は [自己評価点 - 教員評価点] を示している。したがって、上位 2 群 (a と b) は、3 つの評価領域のすべてにおいて、自己評価の方が教員評価よりも低いため、評価値の差は負の値を示している。逆に、下位 2 群 (c と d) では、3 領域のすべてにおいて自己評価の方が教員評価より高いため、評価値の差は正の値を示している。図表 4 の平均点を眺めて最も目を引くのは、すべての評価領域で教員の評価が最も高い b クラスの自己評価値が、他のどのクラスよりも低くなっていることである。教師が最も高く評価した学生が、自分たちのパフォーマンスを最も低く評価していたのである。

これらの差異が偶然のものか、何らかの原因で必然的に生じているものなのかを調べるために、a から d までの 4 群のそれぞれで、自己評価値と教員評価値の差を「対応のある t 検定」で調べた。すると、4 群中 c クラスだけは、どの領域においても教員と学生の評価に有意差がなかった。一方、上位 2 群と最下位群の d クラス



では、a クラスの「態度」でのみ有意差がなかったものの、それ以外のすべての場合において、教員と学生との評価に有意な違いがあった。これは、c クラスの学生だけは、すべての面で、教員の作成したループリック表の基準に従って、適切な自己評価ができたが、その他のクラスの学生は、a クラスの「態度」の評価を除いて、ループリックを用いても教員と同じようには評価できなかったということを示している。

3.1.2 自己評価が教員評価と一致しなかった要因

それでは、上位 2 群 (a と b) の自己評価が教員よりも低かった理由、逆に最下位群の d の自己評価が教員よりも高い理由は何なのだろうか。図表 2 から明らかのように、上位 2 群と下位 2 群で自己評価と教員評価の差が対照的になっていることから、明らかに習熟度の違いが関係していると考えられた。そこで、習熟度の個人差を示す指標として、(1) a-b-c-d 間のクラス差、すなわち、入学時点で測定された習熟度の個人差と、(2) 今回の試験で教員が与えた評価、すなわち、入学後約半年 (a と b の場合は 4 月中～7 月中) ～1 年 (c と d の場合は 4 月中～翌年 1 月中) 後の試験実施時点での発話能力の個人差という 2 つの個人差を用い、この 2 つの指標が自己評価と有意に相関しているのかどうかを評価領域ごとに調べた。(1) を見るためには、自己評価値と教員評価値との差とクラス (a～d の 4 順位) の相関を調べた。また、(2) を見るためには、自己評価値と教員評価値との差と教員評価値の相関を調べた。すると、図表 6.1～6.4 及び図表 7.1～7.4 のような相関図が得られた。また、相関係数は各図表の下の通りである。評価値は、通常の量的数値データとしてではなく、順位データとして分析し、分析結果にはスピアマンのローを順位相関係数として用いた (** $p < 0.01$)。

図表 6.1～6.4 は、クラスが最上位 (1= a クラス) になるほど教員評価値と自己評価値の正の差が大きくなる、すなわち、教員評価よりも自己評価が低くなるとい

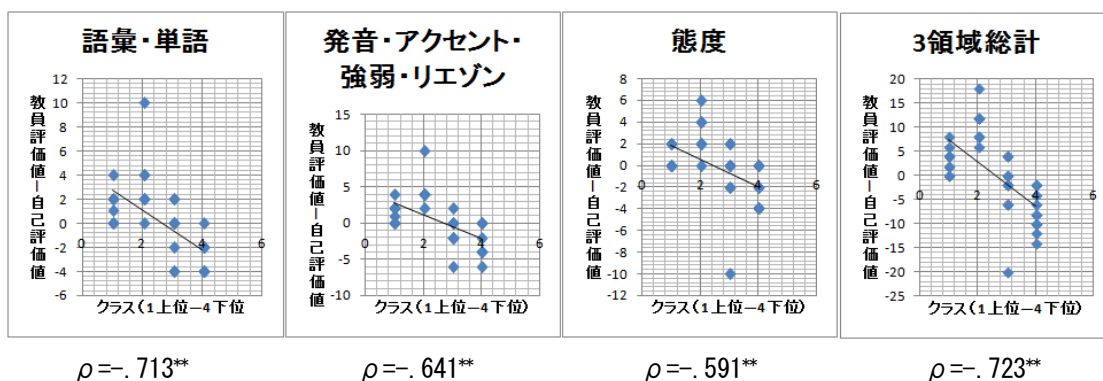
うこと、また、逆に、最下位 (4=dクラス) になるほど負の差が大きくなる、すなわち、教員評価よりも自己評価の方が高くなるということをわかりやすく示している。同様に、図表 7.1~7.4 は、授業の最後の試験での教員評価が高くなればなるほど、教員評価より自己評価が低くなるために教員評価値と自己評価値の正の差が大きくなり、逆に、教員評価が低くなればなるほど、教員評価より自己評価が高くなるために教員評価値と自己評価値の負の差が大きくなることを明確に示している。どの3領域も、中程度の相関~強い相関を示しており、全領域の評価値を総合した「3領域総計」(図表 6.4 及び 7.4) では、有意水準 0.01 レベルでそれぞれ $\rho = -.723$ 、 $\rho = .762$ という強い相関を示した。つまり、習熟度の高い学生ほど自分のパフォーマンスを過少評価し、習熟度の低い学生ほど自分のパフォーマンスを過大評価する傾向のあることが統計的に示されたことになる。

図表 6.1

図表 6.2

図表 6.3

図表 6.4

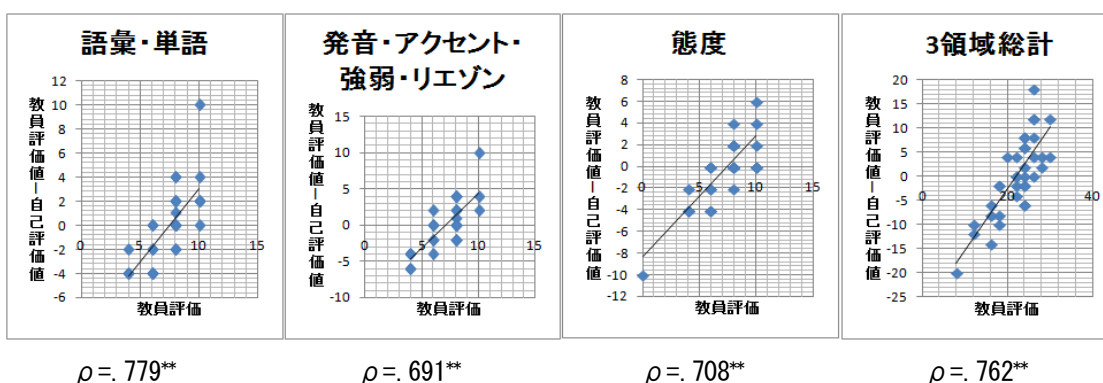


図表 7.1

図表 7.2

図表 7.3

図表 7.4



3.2 各質問に対する解答の集計結果

3.2.1 全4層のアンケート回答の集計方法について

アンケートでは全部で 37 項目の質問を行った。パート 1 で得た上位 2 層 (a と b) と今回の実践研究で得られた下位 2 層 (c と d) のアンケートデータを併せて改めて分析してみると、次のことが分かった(図表 8)。なお、2 群間の分析にはカテゴリカルデータ用の Man-Whitney の U 検定を、4 群間の分析には、Kruskal-Wallis の検定とボンフェローニの修正による多重比較^{*}を行った。

図表 8

対照集団	有意差が検出された項目
上位2層 (aとb)	なし
下位2層 (cとd)	質問4「リスニング試験を含む筆記による試験（以下「筆記試験」）とループリックを用いたインタビュー試験について」の d.「こういう評価表では自分の英語力を正当に評価してもらえない」(p<.05)
上位2層 (aとb) と 下位2層 (cとd)	質問2「試験の前のビデオ撮影について」 a.「自分が英語をどのように話しているか客観的にみることができた」(p<.01) b.「自分の発音と改善すべき点を自分で後から確認することができた」(p<.01) 質問4「筆記試験とループリックを用いたインタビュー試験について」 b.「ループリックを用いたインタビュー試験をするほうが筆記試験をするよりも習った英語がよく身につく」(p<.05) 質問5「ループリックのような評価表について」 d.「こういう評価表では自分の英語力を正当に評価してもらえない」(p<.01) 質問6「ループリックとビデオ撮影を用いた試験について」 d.「英語学習には必要な手段だと思う」(p<.01)
4層* (a-b-c-d)	質問1「試験の前のビデオ撮影について」 a.「自分が英語をどのように話しているか客観的にみることができた」で、上位クラス a、b と最下位クラス d の間に有意傾向 (a-d: p=.009、b-d: p=.036**) b.「自分の発音と改善すべき点を自分で後から確認することができた」で、上位クラス a、b と最下位クラス d の間に有意傾向 (a-d: p=.022、b-d: p=.025**) 質問5「ループリックのような評価表について」の b.「こういう評価表があると、もっと上を目指して頑張ろうと思える」で、上位クラス a、b と最下位クラス d の間に有意傾向 (a-d: p=.022、b-d: p=.025**) d.「こういう評価表では自分の英語力を正当に評価してもらえない」で、上位クラス b と最下位クラス d の間に有意差 (p=.007<.0083*) 質問6「ループリックとビデオ撮影を用いた試験について」の c.「英語の学習の参考にはなったと思うが、英語の上達につながったとは思えない」で、上位クラス a、b と最下位クラス d の間に有意傾向 (a: p=.029、b: p=.023**) d.「英語学習には必要な手段だと思う」で、上位クラス a と下位クラス c、d、上位クラス b と最下位クラス d の間に有意傾向 (a-c: p=.050、a-d: p=.013、b-d: p=.023**)

※ ボンフェローニの修正：4 群をペアにする組合せの数は 6 ペアなので、有意水準 0.05 での 2 群の検定 (Mann-Whitney の検定) を 6 回繰り返すものと考え、帰無仮説の棄却基準を $0.05 \div 6 = 0.0083$ 未満に下げて行った。

以上のように、今回の実践研究の対象となった下位 2 層 (c と d) の間では、質問 4「筆記試験とループリックを用いたインタビュー試験について」の項目 d.「こういう評価表では自分の英語力を正当に評価してもらえない」以外は、有意に異なる回答パターンは見つからなかった。また、上位 2 層 (a と b) の間では、どの質問項目においても、有意に異なる回答パターンは検出されなかった。同時に、上位 2 層をまとめた上位群と下位 2 層をまとめた下位群の間に有意差が検出された 6 つの質問項目と、4 層を一度に比較検定した場合に有意差が認められた 6 つの質問項目のうち 5 つが一致した。

これらのことから、ループリック評価やビデオ撮影に関するいくつかの項目については、習熟度が相対的に高いか低いかによって学生の感じ方に有意な差があることが示されたので、そのような違いが存在する項目を調べた。

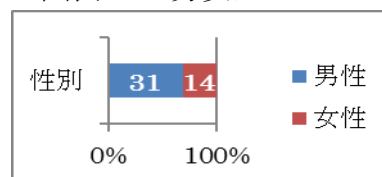
以下の報告では、紙面の制約上、c クラス、d クラスの個々の統計量の報告は割愛し、a、b、c、d4 層のデータを一括して報告する。ただし、上位 2 層・下位 2 層の 2 群間で有意差のあった項目については、別途、詳しく取り上げることとする。

3.2.2 全4層のアンケート回答集計結果

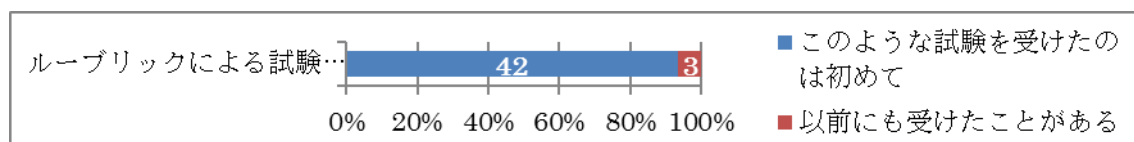
図表9 習熟度別4層の内訳



図表10 男女比



図表11 ルーブリックによる試験について

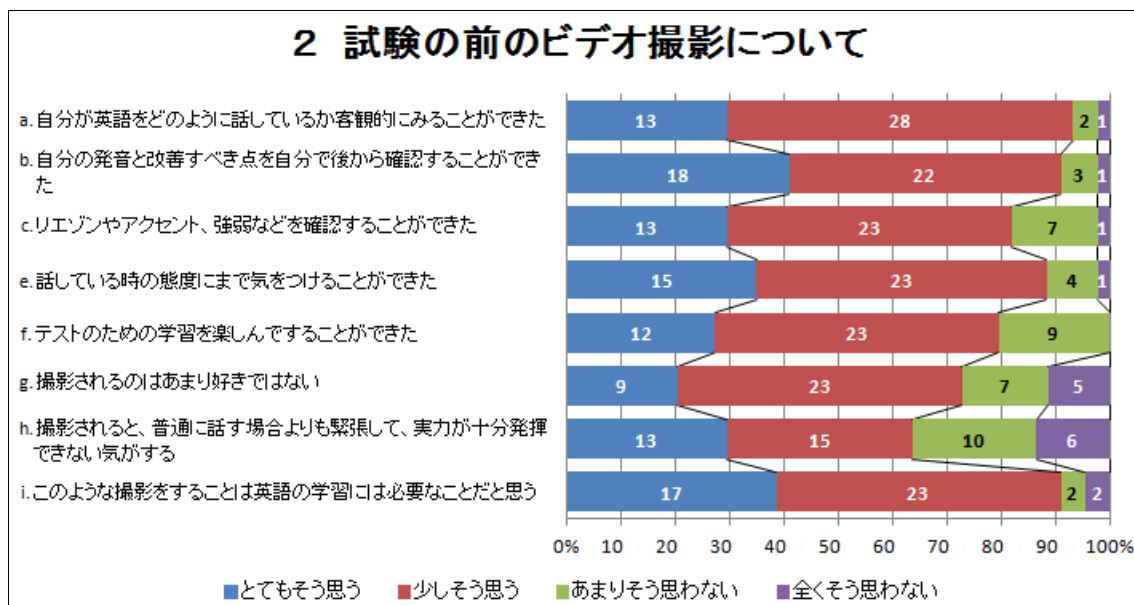


ほとんどの学生(93%)にとってルーブリック評価は初めての体験で、初等・中等教育の英語教育ではまだほとんどルーブリック評価が行われていないということがわかった。

3.3.2.1 質問2「試験前のビデオ撮影について」

図表12の度数分布を見れば明らかのように、スマートフォンによるビデオ撮影は、全体としては有用な手段として肯定的にとらえているが、ビデオ撮影そのものについては、「好きではない」、「撮影されると…緊張して実力が発揮できない気がする」と答えた学生が6~7割に上った。

図表12



(「質問d」が欠けているのは、「d」を削除後「e」以下を繰り上げずにアンケートをとったことによるもので、本報告でもそのまま記載することとした。)

また、先に述べたように、この質問では、項目 a. 「自分が英語をどのように話しているか客観的にみることができた」と項目 b. 「自分の発音と改善すべき点を自分で後から確認することができた」について、上位層(a と b)と下位層(c と d)の2

群間に有意差があった。クロス集計して確認すると図表 13. 1、13. 2 の結果となった。項目 a、b とも、上位群では「とてもそう思う」が有意に多く、下位群では有意に少ない。したがって、上位と下位で差があると言っても、それは、「とても」と答えるか、「少し」と答えるかの差であって、両群とも基本的には肯定的な回答をしていることには変わらない。

図表 13. 1

ただ、ここで注意したいのは、否定的な回答をしたのはすべて下位層の学生で、上位層にはいなかったということである。a. では全体の約8%にあたる4名、b. では全体の全体の約7%にあたる3名が、自分の英語を自分一人の

	a. 自分が英語をどのように話しているか客観的にみる事ができた				合計
	とてもそう思う	少しそう思う	あまりそう思わない	全くそう思わない	
成績上位群(a,b)	13 59.1%(2.5)	9 40.9%(-1.2)	0 0%(-1.8)	0 0%(-1.0)	22 100%
成績下位群(c,d)	5 22.7%(-2.5)	13 59.1%(1.2)	3 13.6%(1.8)	1 2.3%(1.0)	22 100%
合計	18 40.9%	22 50.0%	3 6.8%	1 2.3%	44 100.0%

%の右の()は調整済み残差 (残差が 1.96 より大きくなるか、-1.96 より小さくなれば、有意水準 .05 で有意とみなす)

力でうまくモニターできなかつたと答えた。ビデオを見れば、誰でも自分を客観的に評価できるようになるわけではない、あるいは、その回数が1度では足りないということを示していると言える。

図表 13. 2

	b. 自分の発音と改善すべき点を自分で後から確認することができた				合計
	とてもそう思う	少しそう思う	あまりそう思わない	全くそう思わない	
成績上位群(a,b)	11 50%(3)	11 50%(-1.9)	0 0%(-1.4)	0 0%(-1)	22 100%
成績下位群(c,d)	2 9.1%(-3)	17 77.3%(1.9)	2 9.1%(1.4)	1 4.5%(1)	22 100%
合計	13 29.5%	28 63.6%	2 4.5%	1 2.3%	44 100.0%

%の右の()は調整済み残差

3. 2. 2 質問3「練習時間について」

練習時間については、7割(32人、71%)の学生が「ちょうど良かった」と回答している。「もう少し練習時間が多い方が良かった」と答えた学生が、どれくらいの練習時間があれば「ちょうど良い」と思えるのかは今回の調査では調べられなかった。しかし、質問5の項目e. 「こういう評価表をもっと頻繁に使えば、自分の向上の度

図表 14		5. ルーブリックのような評価表について : e. こういふ評価表をもっと頻繁に使えば、自分の向上の度合いがより分かりやすくなる)			合計
		とてもそう思う	少しそう思う	あまりそう思わない	
3. 練習時間について	ちょうど良かった	6 18.8% (0.3)	23 71.9% (0.7)	3 9.4% (-1.2)	32 100.0%
	もう少し練習時間が多い方が良かった	2 15.4% (-0.3)	8 61.5% (-0.7)	3 23.1% (1.2)	13 100.0%
合計		8 17.8%	31 68.9%	6 13.3%	45 100.0%

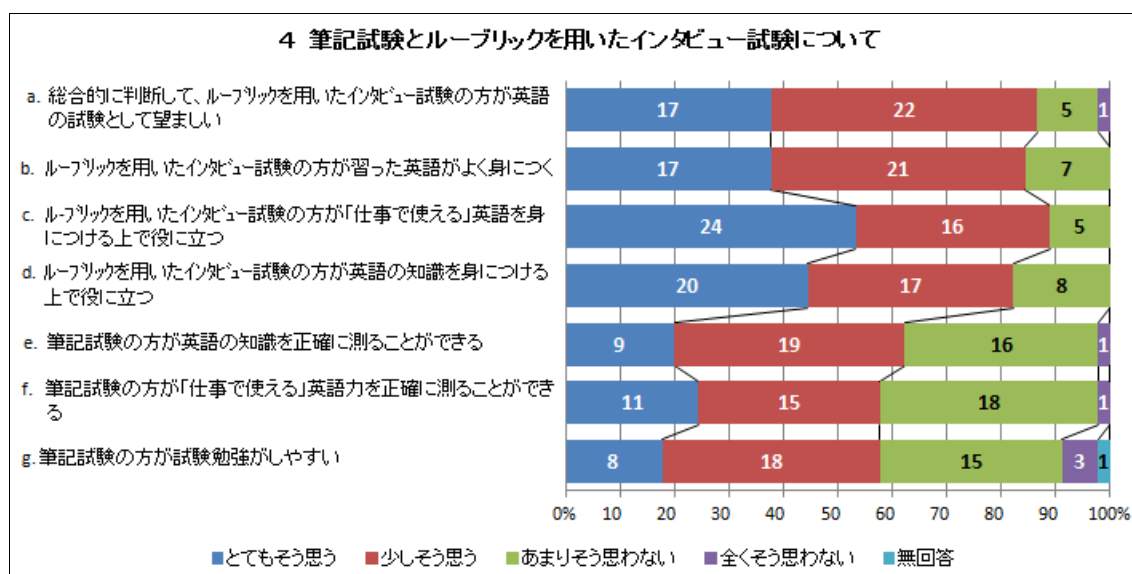
%の右の()は調整済み残差

合いがより分かりやすくなる」とクロス集計をしてみると(図表 14)、質問 3 で回答が分かれた学生たちも、質問 5e では、統計的に有意差のない割合で「とても」または「少し」「そう思う」と答えている。つまり、練習時間がちょうど良いと感じたかどうかにかかわらず、8~9 割の圧倒的多数の学生が、もう少し練習していればもっと役に立ったかもしれないと思っていたと言える。

3.2.3 質問 4「リスニング試験を含む筆記による試験(以下「筆記試験」)とループリックを用いたインタビュー試験について」

図表 15 に示されている通り、筆記試験とループリックを用いた試験を比べた場合、全ての層を総合して見ると、筆記試験よりもループリックを用いた試験の方が『仕事で使える』英語を身につける上で役に立つ(質問 4c)・「習った英語がよく身につく」(質問 4b)で「とてもそう思う」・「少しそう思う」と答える学生が圧倒的多数を占めた。学生たちが、1 度限りの経験でも、ループリックを用いた試験の方を肯定的に受け止めたということがわかった。ループリック評価のもたらす肯定的な影響力の大きさが証明されたと言える。

図表 15



ただ、項目 d の「ループリックを用いた試験をする方が、筆記試験をするよりも英語の知識を身につける上で役立つ」では、下位 2 層の c と d の 2 群間で有意差があった ($p < .05$)。図表 16 にあるように、c クラスの方が「とてもそう思う」の回答

図表 16	4d. ループリックを用いたインタビュー試験をする方が筆記試験をするよりも、英語の知識を身につける上で役に立つ			合計
	とてもそう思う	少しそう思う	あまりそう思わない	
c クラス	8 66.7% (2.3)	3 25% (-1)	1 8.3% (-1.6)	12 100.0%
d クラス	2 18.2% (-2.3)	5 45.5% (1)	4 36.4% (1.6)	11 100.0%
合計	10 17.8%	8 68.9%	5 13.3%	23 100.0%

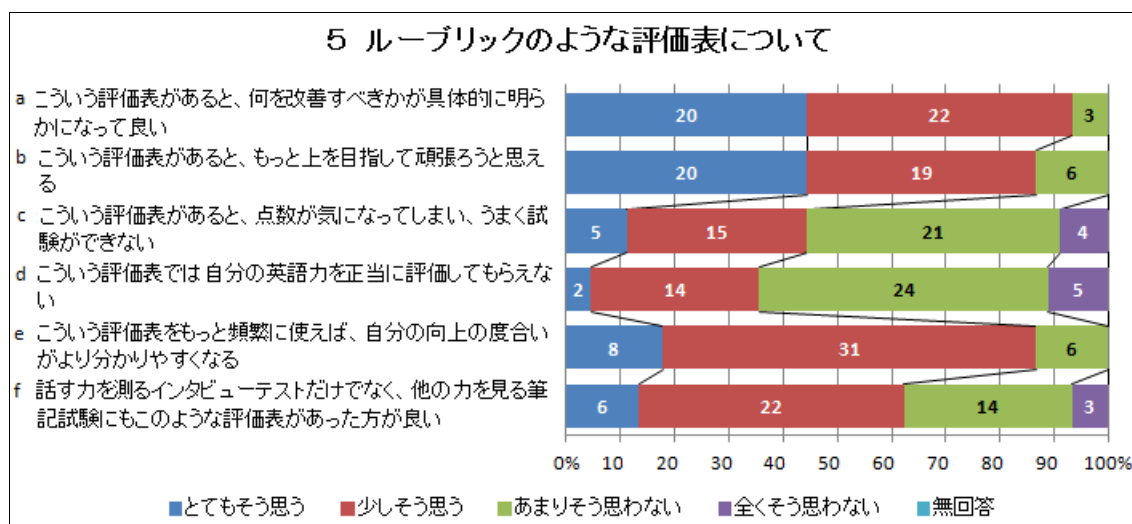
%の右の()は調整済み残差

が有意に多かった。ルーブリックそのものが「英語の知識」獲得に役立つとは考えにくく、最下位層のdクラスの反応は極めて妥当である。その意味で、この結果は、他の3層が筆記試験よりもインタビュー試験の方を肯定的に受け止めていた度合いがいかに強いものであったかを示していると言える。

3.2.4 質問5「ルーブリックのような評価表について」

図表17で、項目a、b、eの回答パターンを見ると、圧倒的多数の学生が「とてもそう思う」と「少しそう思う」を選び、肯定的な反応を示した。これらの回答を見ると、ルーブリックには、学生を自分の課題に気づかせ、自分に合った「次」の目標に向かって自主的に努力し始めるよう動機づける効果があるように見えるが、同時に、質問5cと5dでは、ルーブリック評価の否定的な側面も明確になった。半数近い学生(20人、44.4%)が「点数が気になってしまい、うまく試験ができない」

図表17



と答えた。ただ、上位・下位の2層間の有意差はなかった。また、質問5dへの回答に見られるように、ルーブリック評価では「自分の英語力を正當に評価してもらえない」と感じている学生が(16人、35.6%)もいる。しかし、どんな学生がそう感じているのかと言うと、図表18に示す通り、質問5dでは上位層2群(a、b)と下位層(c、d)との間に有意差のあることがわかった($p < .05$)。ルーブリックでは「自分の英語力を正當に評価してもらえないと思う」と答えた16人の学生のほとんど(13人、

図表18	d こういう評価表では自分の英語力を正當に評価してもらえない				合計
	とてもそう思う	少しそう思う	あまりそう思わない	全くそう思わない	
成績上位群(a、b)	0 0% (-1.4)	3 13.6% (-2.5)	15 68.2% (2)	4 18.2% (1.5)	22 100%
成績下位群(c、d)	2 8.7% (1.4)	11 47.8% (2.5)	9 39.1% (-2)	1 4.3% (-1.5)	23 100%
合計	2 4.4%	14 31.1%	24 53.3%	5 11.1%	45 100.0%

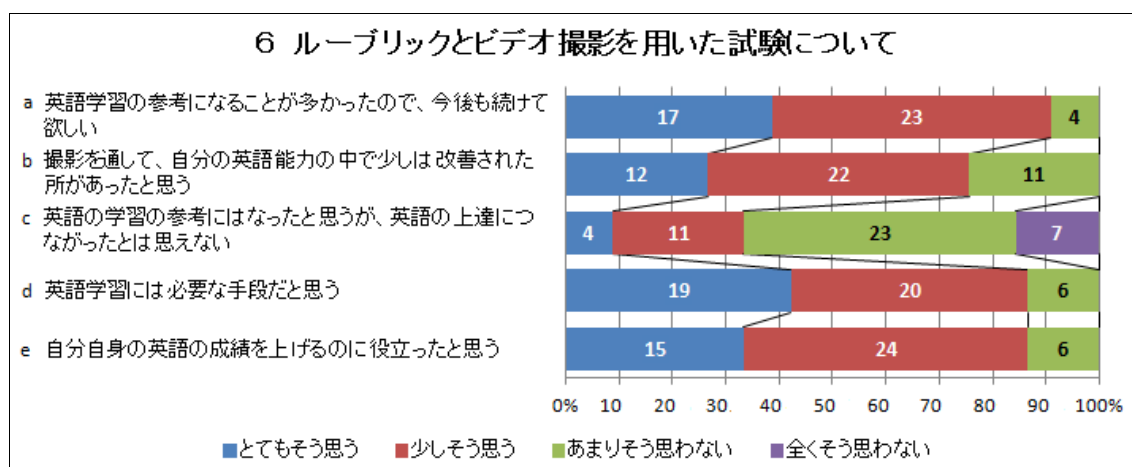
%の右の()は調整済み残差

81%)が下位層の学生だった。これは、ルーブリック表を使って自身のパフォーマンスがどのように点数化されるのかをしっかりと理解するのに、下位層の学生は上位層の学生以上に時間を要する可能性を示唆していると言える。

3.2.5 質問6「ルーブリックとビデオ撮影を用いた試験について」

図表 19 に明らかなように、ルーブリックとビデオ撮影を組み合わせた試験について尋ねた質問6においても、すべての質問項目で、ルーブリックとビデオ撮影を用いた試験は、全ての層において肯定的に受け止められたと言える。

図表 19



しかし、質問6の「d. 英語学習には必要な手段だと思う」で上位層と下位層との間で有意差があった($p < .05$)。図表 20 に示されている通り、上位群には「思わない」と回答をした者が一人もいないのに対して、下位層では、「思わない」と否定的回答をしている者が有意に多くいた。質問5dで見たのと同様、下位層には、このような取り組みをどのように活用するかを丁寧に伝える必要があるということがわかる。

図表 20

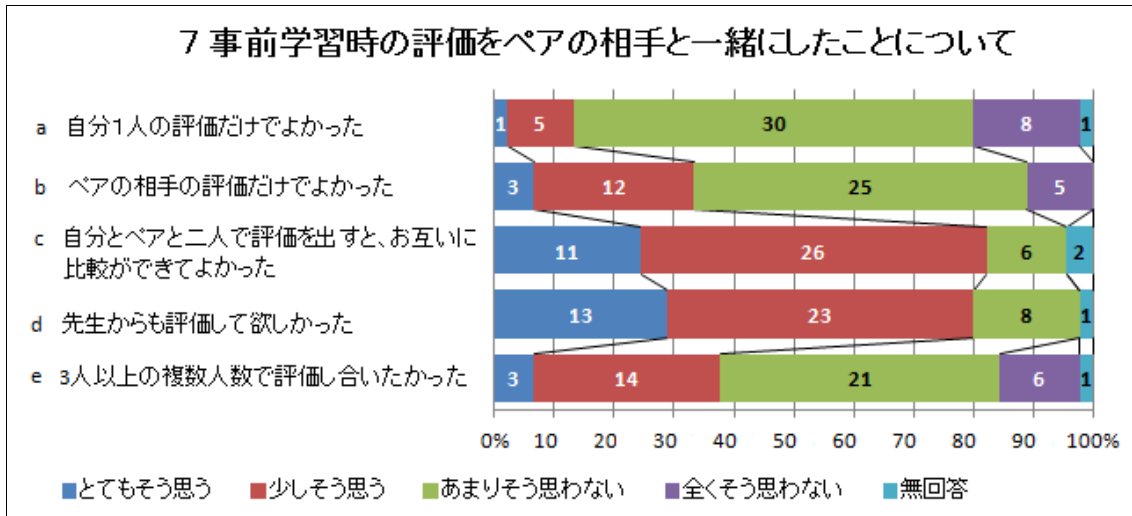
	6d (ルーブリックとビデオ撮影を用いた試験は)英語学習には必要な手段だと思う			合計
	とてもそう思う	少しそう思う	あまりそう思わない	
成績上位群(a,b)	14 63.6%(2.8)	8 36.4%(-1.1)	0 0%(-2.6)	22 100.0%
成績下位群(c,d)	5 21.7%(-2.8)	12 52.2%(1.1)	6 26.1%(2.6)	23 100.0%
合計	19 42.2%	20 44.4%	6 13.3%	45 100.0%

※の右の()は調整済み残差

3.2.6 質問7「事前学習時の評価をペアの相手と一緒にしたことについて」

図表 21 の通り、ほとんどの学生が、ルーブリックを与えられても、ルーブリックに沿ってつけた評価が正しいかどうかを自分一人だけでは判断できないと感じていることがわかった。8割の学生が自分の評価に加えて教員の評価が必要だと感じているということは、学生にとって、自己評価と教員評価を照らし合わせて、自分の評価の正しさを確認する過程が必要であることを示していると言える。その心理は、質問7bと7eの回答パターンにも表れている。

図表 21



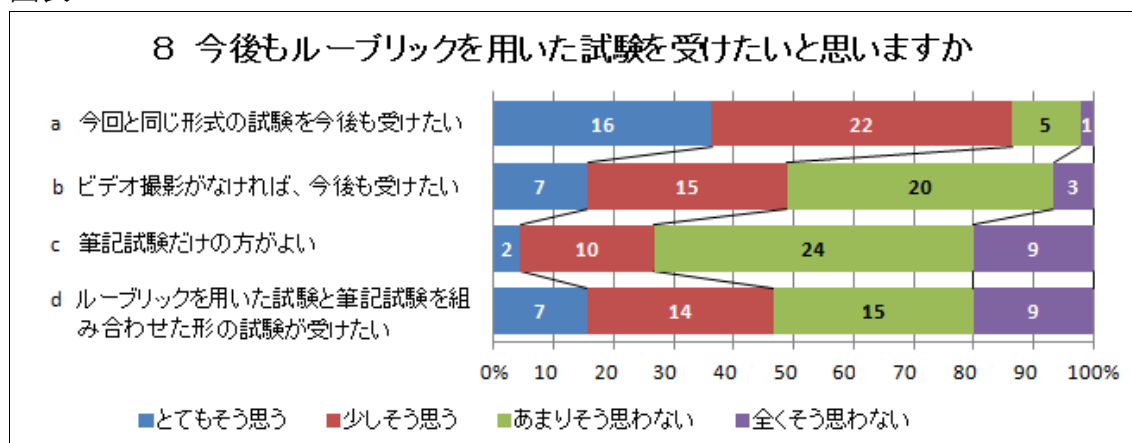
つまり、ルーブリックによる評価の練習は、自分一人では心もとなく（質問 7a）、もう一人のクラスメートの評価も見ながら進めたい（質問 7c）し、教員評価と対照してより正確に評価できるようにもなりたい（質問 7d）が、自分のパフォーマンスの動画をそれ以外の人に見せるということは気恥ずかしい（質問 7e）、と感じている可能性があると考えられる。実際に、質問 7e の回答は、習熟度の差（クラス差）とも、撮影されるのが好きか嫌い（質問 2g）とも有意な相関が見られなかったため、英語で会話するのは好きだが人前での発表は苦手というのと同じ心理が働いたものと考えられる。

3.2.7 質問 8 「今後もルーブリックを用いた試験を受けたいと思いますか」

図表 22 に示されている通り、まず、質問 8a への回答によれば、圧倒的多数の学生がルーブリックによる試験を望んでいることが明確になった。ただ、質問 8b の「ビデオ撮影がなければ、今後も受けたい」という項目への回答は、肯定派と否定派がほぼ半分ずつに分かれた（22 人:23 人）。層別では、a-b-c-d 順に 4:6、5:7、7:5、6:5（人）で、どの層でも肯定派と否定派が拮抗している。この結果が示す意味を理解するために、質問 8b と関連性の深い質問 2g 「撮影されるのはあまり好きではない」、質問 2h 「撮影されると、普通に話す場合よりも緊張して、実力が十分発揮できない気がする」や、質問 5c 「こういう評価表があると、点数が気になってしまい、うまく試験ができない」との相関を見てみた。すると、上位層（a、b 群）では相関が見られなかった（8b-2g、8b-2h、8b-5c それぞれ $\rho = .083$ 、 $.235$ 、 $.200$ ）が、下位層（c、d 群）で高い相関があった（ $\rho = .685^{**}$ 、 $.625^{**}$ 、 $.448^*$ 。 ** は $.01$ 水準で、 * は $.05$ 水準で有意であることを示している）。つまり、下位層では、ビデオ撮影の好悪（2b）やビデオ撮影による緊張（2h）や雑念（5c）といった否定的な要因が、その否定的要因をもたらす「ビデオ撮影（さえ）なければ、今後も受けたい」という気持ちと直接的につながっていたが、上位層ではつながっていなかった。これは、上位層では、「このやり方が良い」という価値判断がビデオ撮影への否定的な感情を相殺するのに対して、下位層では、そのような相殺がなく、単に嫌なものは嫌だという反応に

なっているということを示唆している。つまり、下位層の学生には、今回よりももっと時間をかけてその意義や良さを分からせたり、慣れさせたりしないと、感覚的評価がそのままこの評価方法への信頼度に反映されてしまうということが示唆されている。実際、ビデオ撮影という要因を除いた質問 4g「筆記試験の方が、ルーブリックを用いたインタビュー試験よりも、試験勉強がしやすい」との相関を見ると、上位層では弱い負の相関があった($\rho = -.419$)が、下位層では相関がなかった($\rho = .027$)。これらの結果から、ルーブリックを使った評価・試験を1度限りの経験では、学生はルーブリックの有効性を理解できず、慣れ親しんだ従来型の試験のほうが良いと思ったのではないかと考えられる。よって、このような試験を行う時には、ビデオ撮影の意義やルーブリックの扱い方などの説明はしっかりと時間をかけて行う必要があり、特に習熟の低い層の学生には、なぜ筆記試験ではなくビデオ撮影やルーブリックを試験に用いるのか、どのように活用すべきかを理解してもらう必要があるだろう。

図表 22



5. まとめ

本研究の目標は、冒頭に記述した3つの疑問に答えることであった。その3つに疑問に関しては以下の通りである。

5.1 ルーブリックを用いた評価に対する学生の反応について

パート2のアンケート結果から、ルーブリックを使った評価について学生たちが抱く認識や受け止め方が習熟度によって大きく左右されることはほとんどないということがわかった。今回研究対象とした習熟度の異なる4群間には、4群を比較的上位の2層(a, b)と比較的下位の2層(c, d)に分けた2群の間、及び、上位2層と最下位層(d)の間でいくつかの差異が見られた、自分のパフォーマンスを客観的に見ることができたかどうか(質問 2a)、ルーブリック評価表があると、もっと上を目指して頑張ろうと思えるかどうか(質問 5b)などで、上位層は非常に肯定的なのに対し、下位層は否定的だった。これは、下位層の学生にルーブリックを使うときは、一緒に複数回ルーブリックによる評価練習を重ねながら、その意味や意義を着実に

理解させていく必要があるということの意味している。授業の様子を観察してはじめてわかることなのだが、最下位層のグループには、ビデオ撮影を用いて自身自身のパフォーマンスを客観的に見、評価基準と照らし合わせ、自身の問題点を確認し、向上に役立てるという手順やルーブリックの必要性などを1度限りの練習では理解できない学生が多かった。このことがアンケートの回答に影響したと考えられる。

今回の研究の結論として言えることは、ルーブリック評価は、十分な準備と説明の時間を取ったうえで導入していけば、どのような習熟度の学生にも極めて肯定的に受け入れられ、ルーブリックを通して見えてくる目標に向かって自主的に学んでいこうとする気持ちを生み出す効果的な評価方法の一つだということである。

5.2 ルーブリックを用いた評価そのものについて

パート2では、1度限りのルーブリック評価を練習するだけでも、自己評価と教員評価の差がほとんどなくなり、学生と教員が到達目標の認識を共有することが可能であるかに見える結果が出た。学生の自己評価と教員評価が統計学上有意差が無いとみなせたのは下位層のcグループであった。他の3つのグループでは有意差があるのに、なぜこのグループだけ有意差がなくなったのか。今回集めたデータではその原因を特定することはできなかった。ただ、他の3群では、自己評価と教員評価の間に一致がなかったことは確かである。ルーブリックという共通基準を活用して、学生が教員と同じ目で自分を評価し、目標の達成を確認して、次の目標を自主的に設定できるようにするためには、複数回にわたってルーブリック評価を使った学習をさせ、自己評価の仕方と教師の評価を比べる時間を設け、学生が自身で適切な評価ができるように導いていく必要があると言える。

また、今回の研究では、習熟度が高い学生ほど自分に対する評価が厳しくなり、自分のパフォーマンスを過小評価する一方、習熟度が低い学生ほど自分に対する評価が甘くなり、自分のパフォーマンスを過大評価する傾向のあることが統計的に示された。つまり、未熟あるいは能力の低い個人ほど自分を実際よりも高く評価してしまう認知バイアスとして心理学で広く知られているダニング＝クルーガー効果(Dunning-Kruger effect)が、大学生の英語学習における自己評価の過程においても働いているということがわかった。実際、初等教育レベルの英語教育でも、「学年が上がるにつれて、(1)自尊心が落ちてくる…、(2) 正確さを重視する」という研究や、「学年が上がるにつれて…自己評価における否定的意識が高」くなり、「教師との評価のズレが出てくる」ため、「自己効力感を与えるような形で自己評価をさせないと、下向きに落ち込んで行く」という(久野、2012、p.6)。今回の実践研究で、この傾向は高校卒業後間もない大学生においても見られるということが示唆された。ルーブリック評価そのものは、パフォーマンス評価として様々な利点を持っているが、自己評価の根底にある心理作用に注意しないと、ルーブリック評価への学習者の信頼や学習意欲そのものを削いでしまう恐れがあるということがわかった。

5.3 今回の研究の限界と今後の研究

今回のアンケート調査は、習熟度の下位層と上位層の違いを見るために、パート1と全く同じ質問項目を使うことにした。今後は、質問項目を改良して、学生がビ

デオ撮影を好まないのは、撮影に対する慣れで改善できるものなのか、ビデオを用いた練習をどの程度すれば抵抗感がなくなるのか、試験前の練習時間はどの位が妥当なのかなどについて研究していきたい。

また、ルーブリックによる評価をより効果的なものにするには、ルーブリック評価を複数回にわたり繰り返し練習させながら、その中で教師のルーブリック評価を学生に提示して、学生自身の評価と教師の評価の違いを確認することで、自身のどの部分を改善すればより英語能力の改善につながるのかに気づかせる、すなわち、キャリブレーションを普段の授業に取り入れる必要がある、ということだ。今回の研究では、上位と下位の違いを検証するためにあえてキャリブレーションの取り組みをしなかったが、今後は、キャリブレーションを実際の授業の中に取り入れ、その成果を検証していきたい。

謝辞

最後に、学生のアンケートの集計結果とインタビューテストの評価データの統計的処理を快くお引き受けくださった札幌大谷大学社会学部の久野寛之教授に、この場をお借りして心から感謝の意を述べさせていただきたい。

引用文献

- 當作靖彦(2014) 「学習を促進する評価のデザイン」 配布資料. 国際フォーラム・SPELT・REX・NET 主催. 10月11日札幌市、札幌大谷大学
- 當作靖彦・中野佳代子(2013) 『外国語学習のめやす』 国際文化フォーラム
- 久野寛之(2012) 「小学校外国語活動のこれから—小学校英語教育学会千葉大会の報告」 実用英語教育学会第1回研究会配布資料。10月13日札幌市、札幌大谷大学

付録

今回曜日英語では、ルーブリック評価による対面式インタビューテストを導入しました。今回のテストを今後の評価方法改善に役立てるために、是非みなさんの感想と意見を聞かせてください。記名はしますが、このアンケートが成績に影響することは一切ありません。あくまでも改善目的ですので、思ったことをありのまま答えてください。よろしく願います。

学籍番号 _____ 氏名 _____

該当するものには、() に○をつけてください。

1 ルーブリックを使用した試験について

a () このような試験を受けたのは初めて b () 以前にも受けたことがある

2 試験の前のビデオ撮影について

- a 自分が英語がどのように話しているか客観的にみることができた
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- b 自分の発音と改善すべき点を自分で後から確認することができた
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- c リエゾンやアクセント、強弱などを確認することができた
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- e 話している時の態度にまで気をつけることができた
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- f テストのための学習を楽しんですることができた
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- g 撮影されるのはあまり好きではない
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- h 撮影されると、普通に話す場合よりも緊張して、実力が十分発揮できない気がする
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- i このような撮影をすることは英語の学習には必要なことだと思う
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない

3 練習時間について（事前に1時間の練習をしたことについて）

a () ちょうど良かった b () もう少し練習時間が多いほうが良かった

4 リスニング試験を含む筆記による試験(以下「筆記試験」とルーブリックを用いたインタビュー試験について

- a 総合的に判断して、ルーブリックを用いたインタビュー試験の方が筆記試験をするよりも英語の試験として望ましい
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- b ルーブリックを用いたインタビュー試験をする方が筆記試験をするよりも習った英語がよく身につく
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- c ルーブリックを用いたインタビュー試験をする方が筆記試験をするよりも、「仕事で使える」英語を身につける上で役に立つ
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- d ルーブリックを用いたインタビュー試験をする方が筆記試験をするよりも、英語の知識を身につける上で役に立つ
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- e 筆記試験の方が、ルーブリックを用いたインタビュー試験よりも英語の知識を正確に測ることができる
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- f 筆記試験の方が、ルーブリックを用いたインタビュー試験よりも「仕事で使える」英語力を正確に測ることができる
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- g 筆記試験の方が、ルーブリックを用いたインタビュー試験よりも、試験勉強がしやすい
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない

5 ルーブリックのような評価表について

- a こういう評価表があると、何を改善すべきかが具体的に明らかになって良い
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- b こういう評価表があると、もっと上を目指して頑張ろうと思える
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない

- c こういう評価表があると、点数が気になってしまい、うまく試験ができない
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- d こういう評価表では自分の英語力を正当に評価してもらえない
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- e こういう評価表をもっと頻繁に使えば、自分の向上の度合いがより分かりやすくなる
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- f 話す力を見るインタビューテストだけでなく、他の力を見る筆記試験にもこのような評価表があった方が良い
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない

6 ルーブリックとビデオ撮影を用いた試験について

- a 英語学習の参考になることが多かったので、今後も続けて欲しい
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- b 撮影を通して、自分の英語能力の中で少しは改善された所があったと思う
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- c 英語の学習の参考にはなったと思うが、英語の上達につながったとは思えない
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- d 英語学習には必要な手段だと思う
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- e 自分自身の英語の成績を上げるのに役立ったと思う
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない

7 事前学習時の評価をペアの相手と一緒にしたことについて

- a 自分1人の評価だけでよかった
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- b ペアの相手の評価だけでよかった
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- c 自分とペアと二人で評価を出すと、お互いに比較ができてよかった
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- d 先生からも評価して欲しかった
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- e 3人以上の複数人数で評価し合いがあった
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない

8 今後もルーブリックを用いた試験を受けたいと思いますか

- a 今回と同じ形式の試験を今後も受けたい
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- b ビデオ撮影がなければ、今後も受けたい
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- c 筆記試験だけのほうがよい
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない
- d ルーブリックを用いた試験と筆記試験を組み合わせた形の試験を受けたい
() とてもそう思う () 少しそう思う () あまりそう思わない () 全くそう思わない

9 最後に、今回のルーブリック評価によるインタビュー形式試験を受けてみて、良かった点や改善すべき点（こうしてもらえればもっと良かったと思うことなど）があったら、自由に書いてください。

ご協力、ありがとうございました。

SPELT JOURNAL 投稿規定

1. 執筆者は会員であることを原則とする。
2. 投稿論文は、未発表のものに限る。ただし、口頭発表したものでも、その旨を明記してあれば、審査の対象とする。
3. 投稿分野は、学術的な実験・調査および理論的考察等をまとめた「研究論文」と、教育実践にもとづく知見を報告する「実践研究」との2部に分ける。
4. 原稿は、原則としてMSワードを用い、A4縦長の用紙に以下の書式で作成すること。
 - (ア) 上下左右に3cmずつ余白を設ける。
 - (イ) 横書きで、文字の大きさは、和文・英文とも12ポイントで作成する。
 - (ウ) 和文の場合37字40行でフォントは明朝体、英文の場合74字40行で、フォントはCenturyを使用する。日本語に英数字が混じる場合、全て半角を使用する。
 - (エ) 最初の頁に、3行空けて論文タイトル（16ポイント、ボールドは不要）、1行空けて氏名、改行して括弧書きで所属（共著の場合は改行せずにまとめて可）、さらに1行空けてAbstractと書き、次の行から300語程度の英文Abstractを一つの段落にまとめて置く。論文タイトル、氏名と所属、Abstractの語句はセンタリングを施すこと。Abstractの後、1行空けて本文に入る。
 - (オ) 英文タイトルは、語頭を大文字、他は小文字で表記する。英文著者名はHanako HOKKAIの要領で表記すること。
 - (カ) 各章・節のタイトルには番号をふる。章タイトルはセンタリングし、前後1行ずつ空ける。節タイトルは左寄せし、前後で行を空けない。番号は、ローマ数字ではなく、アラビア数字を用いる。章タイトルの前に付す番号の直後にはピリオドをふる（例：3. 調査の方法）。また、章と節を示す番号の間には、ハイフンの代わりにピリオドを用いる。例：2.1.3
 - (キ) ページ番号は不要。ただし、査読用に印刷したものには、用紙右上に鉛筆で記入する。
 - (ク) Abstract、本文、注、文献リスト、図、表等の資料すべてを含め、研究論文の場合10枚以上20枚以内、実践研究の部の場合5枚以上10枚以内とする。
 - (ケ) 図、表には一連の番号をつける。脚注はつけず、全ての注は本文と文献リストの間にまとめておく。また、表の前後は一行ずつ空ける。
 - (コ) 英文原稿、英文Abstractは、必ずネイティブチェックを受けたものを提出する。
 - (サ) 文献リストは、本文中に言及あるもののみを「引用文献」（日本語論文）あるいはReferences（英語論文）と左寄せして掲げた後に続けて書く。配列は、英語文献を先に著者名のアルファベット順に並べ、次に日本語文献を「あいうえお順」で続ける。論文中の引用の方法や引用文献の書式、図や表の作成方法については、American Psychological Association発行の*APA Publication Manual (5th Ed.)*を参考にすること。参考までに、文献リストの書式例を以下に掲げる。なお、
 - ①日本語文献の発行年のあとのピリオドは不要。

② 紀要等の号数を表示する際、「第～号」のようにする必要はなく、数字だけでよい。

③ インターネット上の資料を挙げる場合には、閲覧日又は取得日を（～年～月～日アクセス：[URL]）又は（～年～月～日現在：[URL]）のように必ず記載する。

Baker, S. C. & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language Learning*, 50, 311-341.

伊田勝憲 (2002) 「学習動機づけの統合的理解に向けて」. 『名古屋大学大学院教育発達科学研究科紀要』, 49, 65-76.

Chamot, A. U. (2001). The role of learning strategies in second language acquisition. In M. P. Breen (Ed.), *Learner contributions to language learning* (pp.25-43). Harlow: Pearson Education Limited.

池田 央 (1989) 「階層クラスター分析」. 池田央(編)『統計ガイドブック』(pp.199-200) 東京：新曜社.

5. 論文本体の構成は以下を基本とする。これ以外の構成の論文投稿を妨げるものではないが、その構成が妥当か否かの判断は査読者の判断に委ねられる。

(ア) 調査・実験等によるデータ収集を伴う研究論文

- ① 「はじめに」または「序論」として、研究の目的及び意義を述べる。
- ② 「研究の背景」や「先行研究のまとめ」等として、当該分野におけるこれまでの関連する研究を概観し、研究課題の新規性・独創性を述べる。（①と②は1つの章にまとめてもよい。）
- ③ 「調査の方法」や「研究の方法」等として、研究課題・仮説（②の最後に述べるか、独立した章を設けてもよい）、調査参加者・実験被験者、データ収集に用いた試験・質問紙等の道具、調査・実験の手続き、教育介入の内容と手順、分析の方法等を述べる。
- ④ 「結果と考察」や「結果と分析」等として、得られたデータのまとめ、その解釈、先行研究で得られた知見との比較等を述べる。「結果」と「考察／分析」の2章に分けてもよい。
- ⑤ 「おわりに」、「むすび」、「結論」等として、新しい知見の概要、教育への示唆、結果解釈上の留意点、今後の研究課題を述べる。

(イ) 理論的考察を行う研究論文 (ア) の論文構成のうち③と④の代わりに、内容に応じて適宜章のタイトルを設け、新しい概念や研究の方向・方法の提示、複数の理論・方法の比較、入手可能な知見による妥当性の考察、当該研究分野に与える影響の説明等を議論する。

(ウ) 教育実践に基づく知見を報告する教育論文

- ① 「はじめに」または「序論」として、取り組みの動機・目的、普遍的問題としての一般化の可能性等を述べる。
- ② 「問題の所在」や「解決すべき問題点」等として、取り組んだ教育実践上の問題点を、深刻さ、解決すべき理由、原因の考察等を述べる。ただし、

- ①と②は1つの章にまとめてよい。
- ③ 「実践の内容」や「取り組み」等として、授業計画、教材、教授の手順、評価方法等を具体的に示しながら、実践の内容を説明する。
- ④ 「効果」や「結果と解釈」等として、観察、評価資料、質問紙・面接・授業評価等の学習者からのフィードバック、第三者の評価等をもとに問題点がどのように、どの程度解決したか、新たな問題は生じなかったか、考えられる理由は何かなどについて説明する。
- ⑤ 「おわりに」、「むすび」、「結論」等として、報告した教育実践の概要と今後の取組の展望を述べる。
6. 原稿を以上の要領で作成し、MSワードの添付ファイルで下記電子メールアドレスまで、9月30日必着で送付すること。（MSワードが使用できない場合は、電子メールにて個別に相談に応ずる。）なお、メールの件名には、*SPELT JOURNAL* 原稿と書き、メール本文に「研究論文」か「実践研究」かの別と、執筆者、所属、連絡先住所を明記すること。氏名・所属を伏せた原稿を下記編集委員長が作成し、査読者に送付する。
- speltjournal@spelt.main.jp
柴田晶子（*SPELT JOURNAL*編集委員長）
札幌大谷大学社会学部地域社会学科 電話：011-742-1651 内線1970
7. 原稿の採否は、査読後決定する。原稿が採用され、掲載される場合には、請求に応じて掲載費（1編5,000円）を支払うこと。
8. 審査後、修正を指示した上で採用する場合、指定された期日まで（通常1週間後）に印刷用原稿を再提出すること。応募段階で届け出た執筆者を採用通知後に変更することは認めないので注意すること。
9. 抜刷りは、第2号以降の電子ジャーナル化に伴い、現在発行していない。

(2015年12月1日改訂)

学会誌委員

柴田 晶子

久野 寛之

三浦 寛子

査読協力者

石川 希美

久野 寛之

柴田 晶子

杉浦 理恵

釣 晴彦

三浦 寛子

『実用英語教育学会紀要』 *SPELT JOURNAL*

ISSN 2187-4123

2016年12月31日 発行

編集者 実用英語教育学会学会誌事務局

発行者 実用英語教育学会

事務局 〒065-8567

札幌市東区北16条東9丁目1番1号

札幌大谷大学社会学部地域社会学科

石川希美 研究室内

電話: 011-742-1651 (代) 内線 1969

FAX: 011-742-1654

E-mail: info@spelt.main.jp

URL: <http://spelt.main.jp>