小泉利恵科研プロジェクト第2回例会 (共催：日本言語テスト学会)

順天堂

JLTA

2021年2月14日（日）　午前9時30分～午後4時10分

テーマ】英語スピーキング評価における採点方法の改善（Part 2）

【主催】小泉利恵科研プロジェクト（基盤研究(C) 20K00894：「高校の英語授業内スピーキング評価における「信頼性確保のための採点指針」の作成」）

趣旨説明 9:35～9:45

- 小泉利恵科研プロジェクト（基盤研究(C) 20K00894：「高校の英語授業内スピーキング評価における「信頼性確保のための採点指針」の作成」）
- 【概要】スピーキング評価 (speaking assessment: SA) において、**タスク・ルーブリック・採点**は特に考慮すべき要素です。本例会では、**採点者や採点トレーニングなどの採点方法**に焦点を当て、**国内の教室内評価や大規模テスト**を含む様々な文脈においてスピーキング評価に携わる方から**運用や課題**についてお話を伺い、今後の日本のスピーキング評価について議論します。

スピーキング評価に関わるもの

Eckes (2019) 訳

タスク
↓
生徒
↓
パフォーマンス ⇒⇒⇒ 教員（採点者）の採点 ⇒⇒⇒ テスト得点 ⇒テスト結果

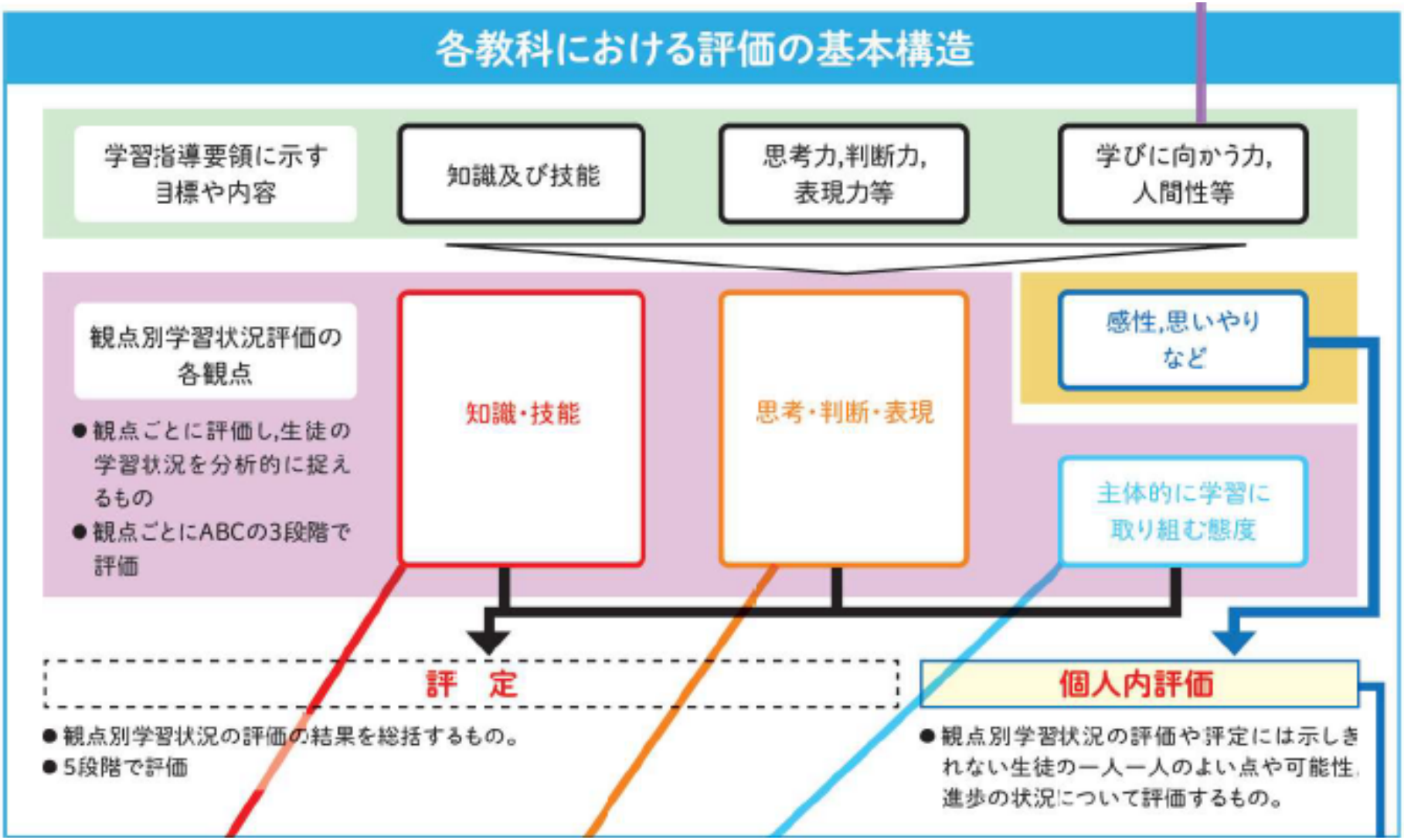
↑
採点基準（ルーブリック）

採点者の影響

- テストが満たすべき3要素 (小泉・印南・深澤, 2017)
- 1. 妥当性：テストが十分適切に、解釈できるか。使えるか
  - 例：指導目標との一致
- 2. 信頼性：テスト結果が十分に安定しているか
  - 大規模テストと違い、授業内テストでは高い信頼性はなくてもよい。中程度は必要
  - 例：採点者間で同じように採点しているか
  - 同じ採点者が同じように採点しているか
  - 同じタスクを行ったときに同じような結果が出るか
- 3. 実行可能性 (実用性)：テスト作成・実施・採点・使用が無理なく行えるか

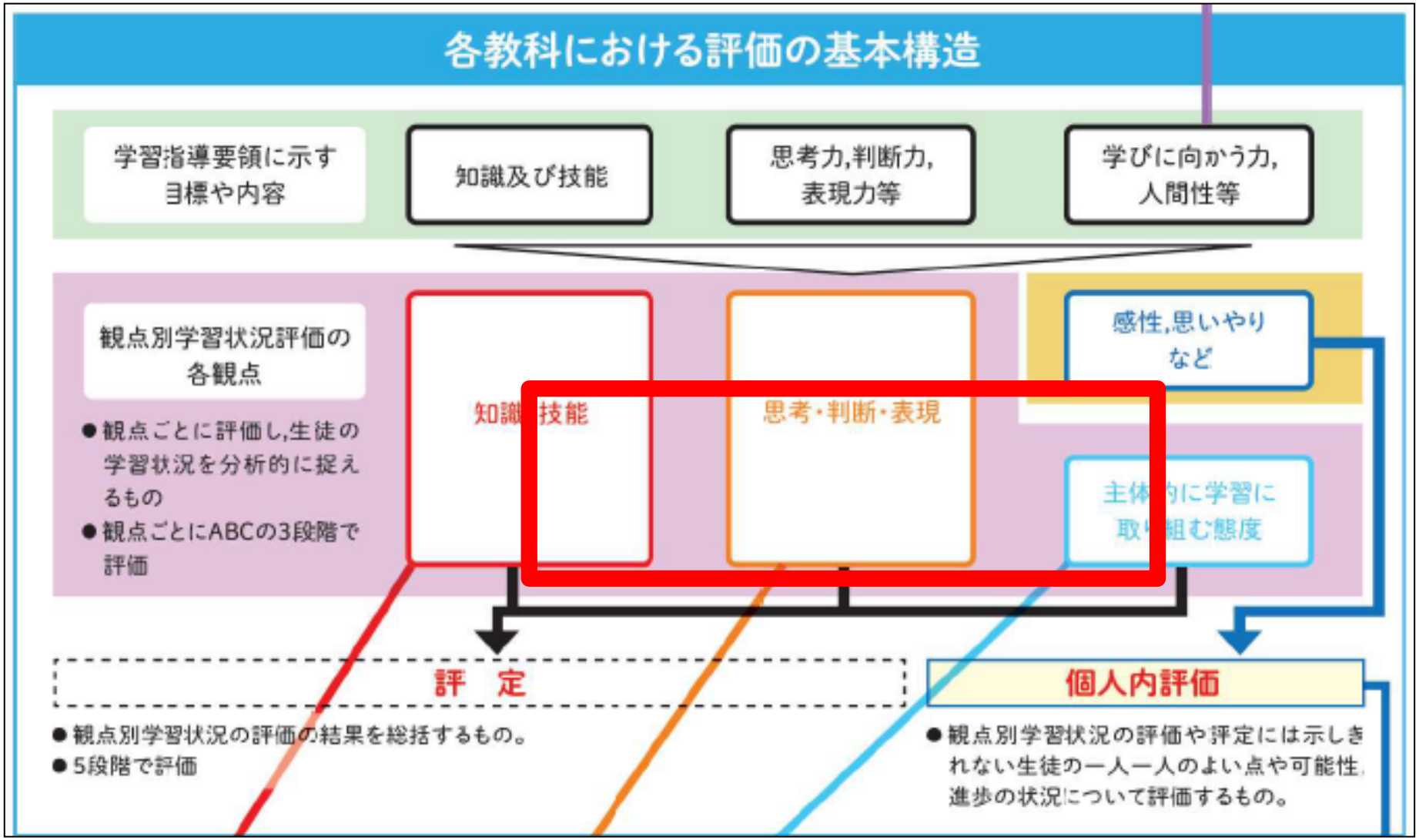
学習評価の方向性

学習評価の在り方ハンドブック（高等学校編）



学習評価の方向性

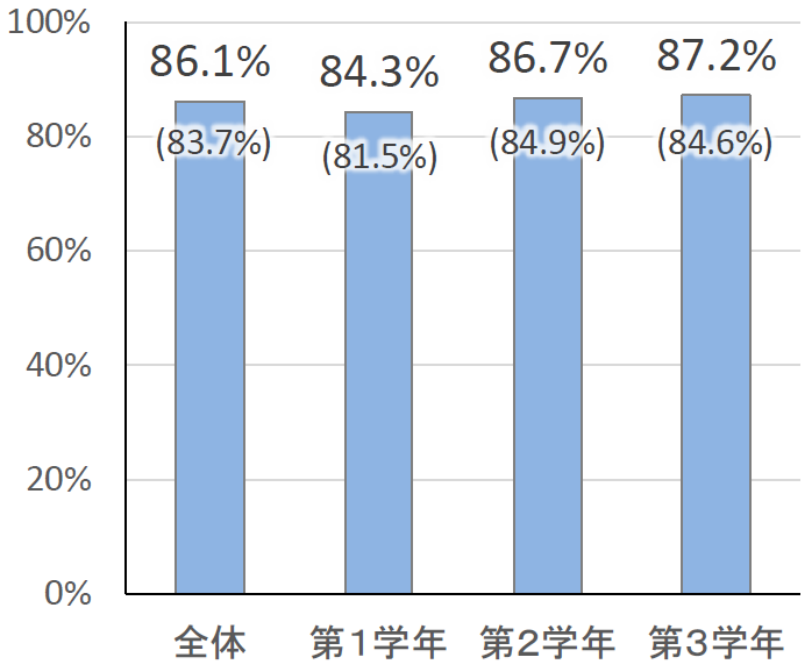
学習評価の在り方ハンドブック（高等学校編）



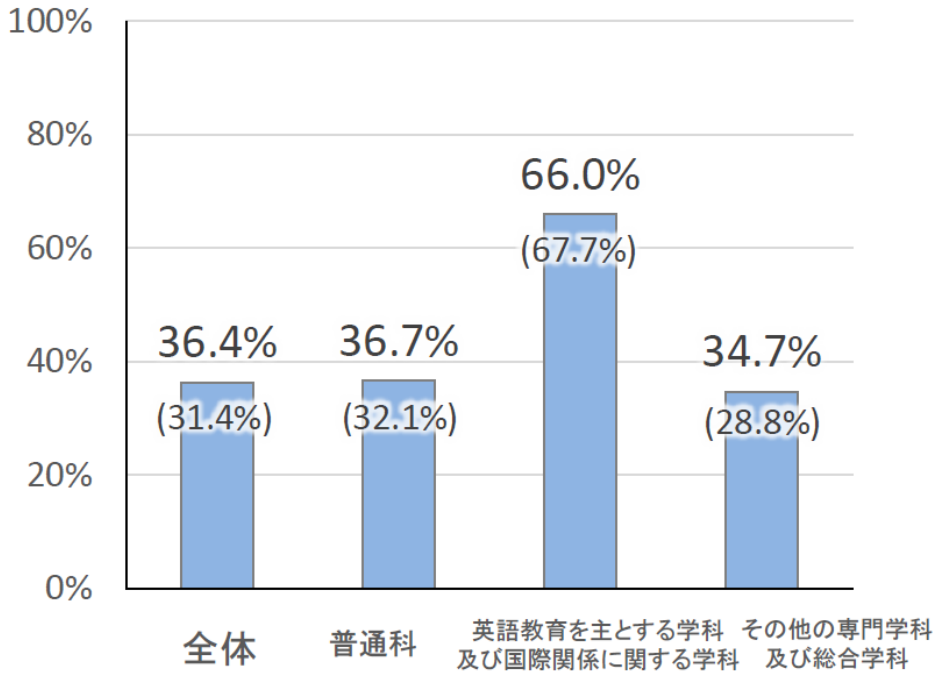
令和元年度「英語教育実施状況調査」の結果

パフォーマンステストの実施状況（中学校・高等学校）

【中学校】

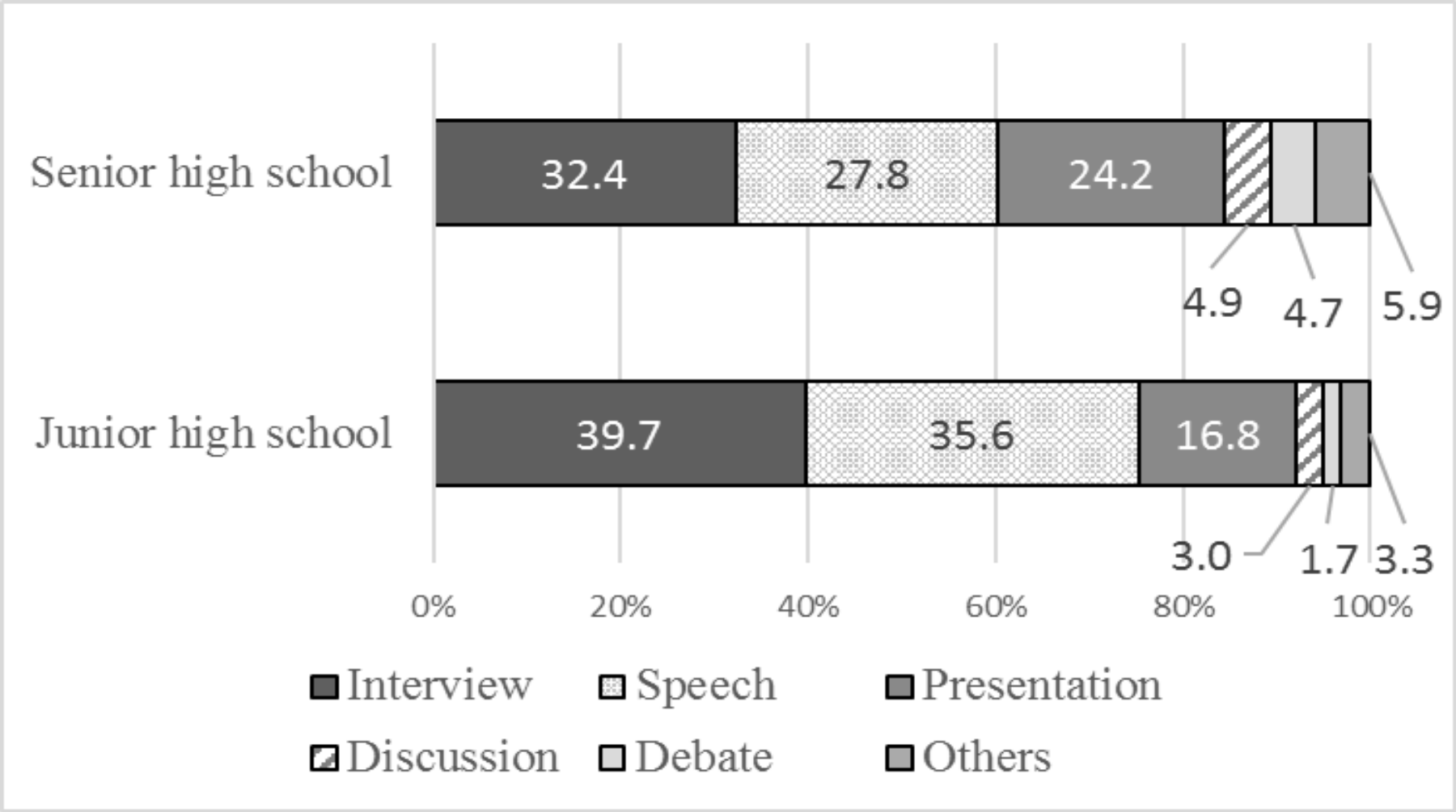


【高等学校】



※（　）内は、昨年度の数値。

使用されるタスク形式
面接型・発表型が多い (文部科学省, 2020に基づく)



さまざまなタスク形式の組みあわせが重要

- 発表型：プレゼン、スピーチ、教科書内容の再生、要約、意見表明（技能統合型）
- やり取り型：面接型、ペア型、グループ型

- 採点では、**ルーブリック（評価基準）**を使用
- 測りたい力を明記し、教師と生徒に理解しやすいものにする。事前に示す

- テストの**採点結果**が十分に**安定**するような方策をとる

タスク例：グループ型会話

4名でのディスカッションでのルーブリック例 (宇都宮南高校)

（1）内容 （テーマ・話し合いの流れに沿った発言内容） 10 点満点

テーマに即した内容について、自分自身の考えや意見についてわかりやすく話している。	10
テーマに即した内容について、流れに沿った発言をして話し合いに参加しているが、内容がやや限定的、または不明瞭である。	7
テーマに即して話そうとしているが、内容がやや不明瞭である。	4

（2）やり取りの適切さ（バランスの取れた発言と他者の意見を聴く姿勢） 10 点満点

他者の意見に耳を傾けたり、既習の会話表現を積極的に用いて反応したり質問したりすることで、グループ内の会話の継続に努めている 。	10
他者の意見にも耳を傾けながら、それについて自分の意見を述べようと努めている。	7
積極性には欠けるが、時々発言したり、他者の意見を聞いたりしている。	4

（3）コミュニケーションへの意欲 （準備した紙を見ずに英語で話そうとする姿勢） 5 点満点

英語のみでやりとりしようとしている。準備した紙をほとんど見ていない。	5
英語でやりとりしようと努めている。準備した紙をたまに見ている。	3
英語での発言が少ない。準備した紙を常に見ている。	1

プロジェクト報告1:
高校の教室内スピーキングテストにおける採点に関する課題

順天堂

JLTA

小泉利恵科研プロジェクト第2回例会

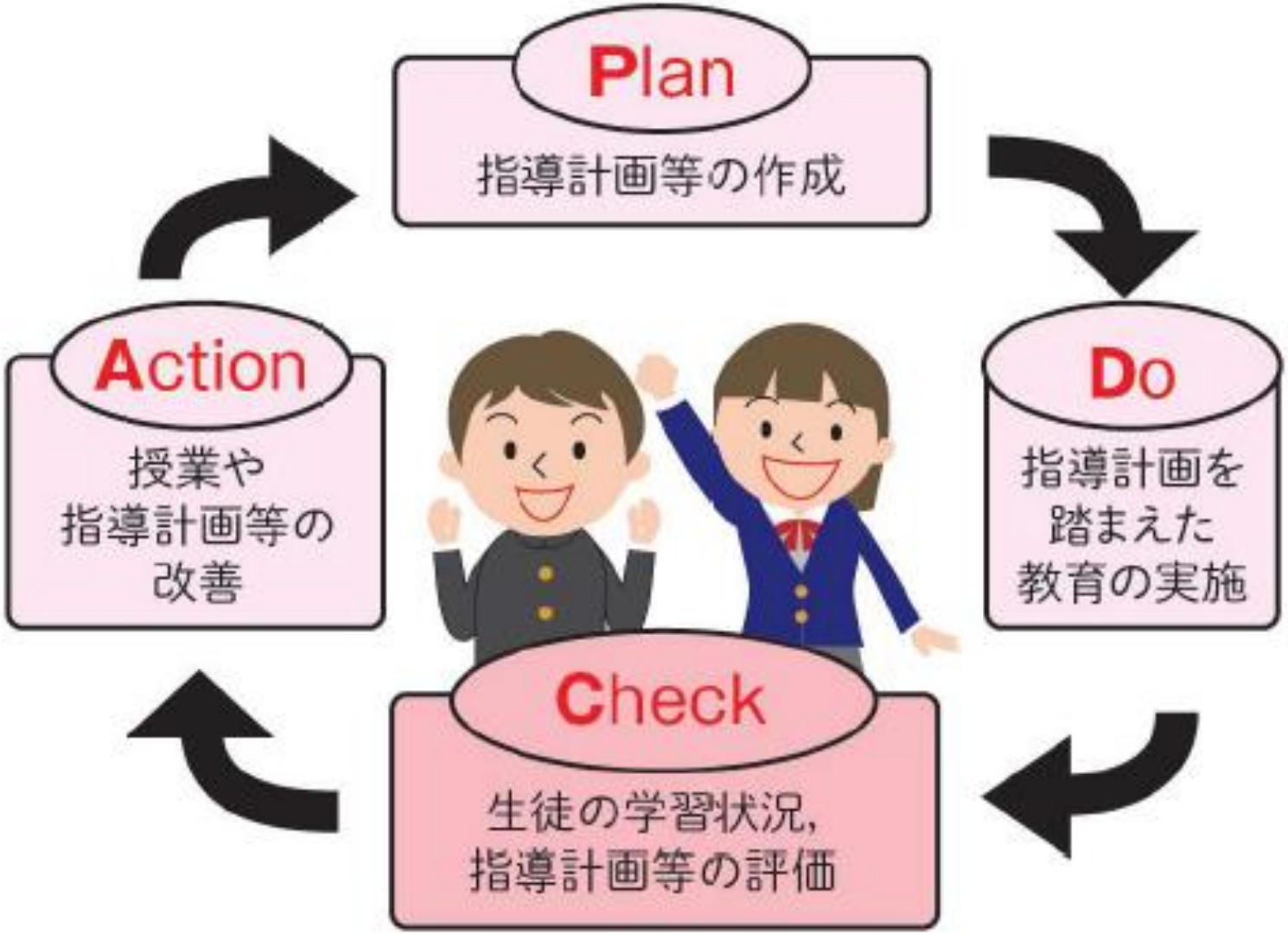
(共催:日本言語テスト学会)

2021年2月14日(日)　14:40～15:10

小泉利恵(順天堂大学)・渡邉聡代(栃木県立宇都宮南高校)・初澤晋(宮城県石巻高校)・磯部礼奈(宮城県石巻高校)・松岡京一(宮城県石巻高校)

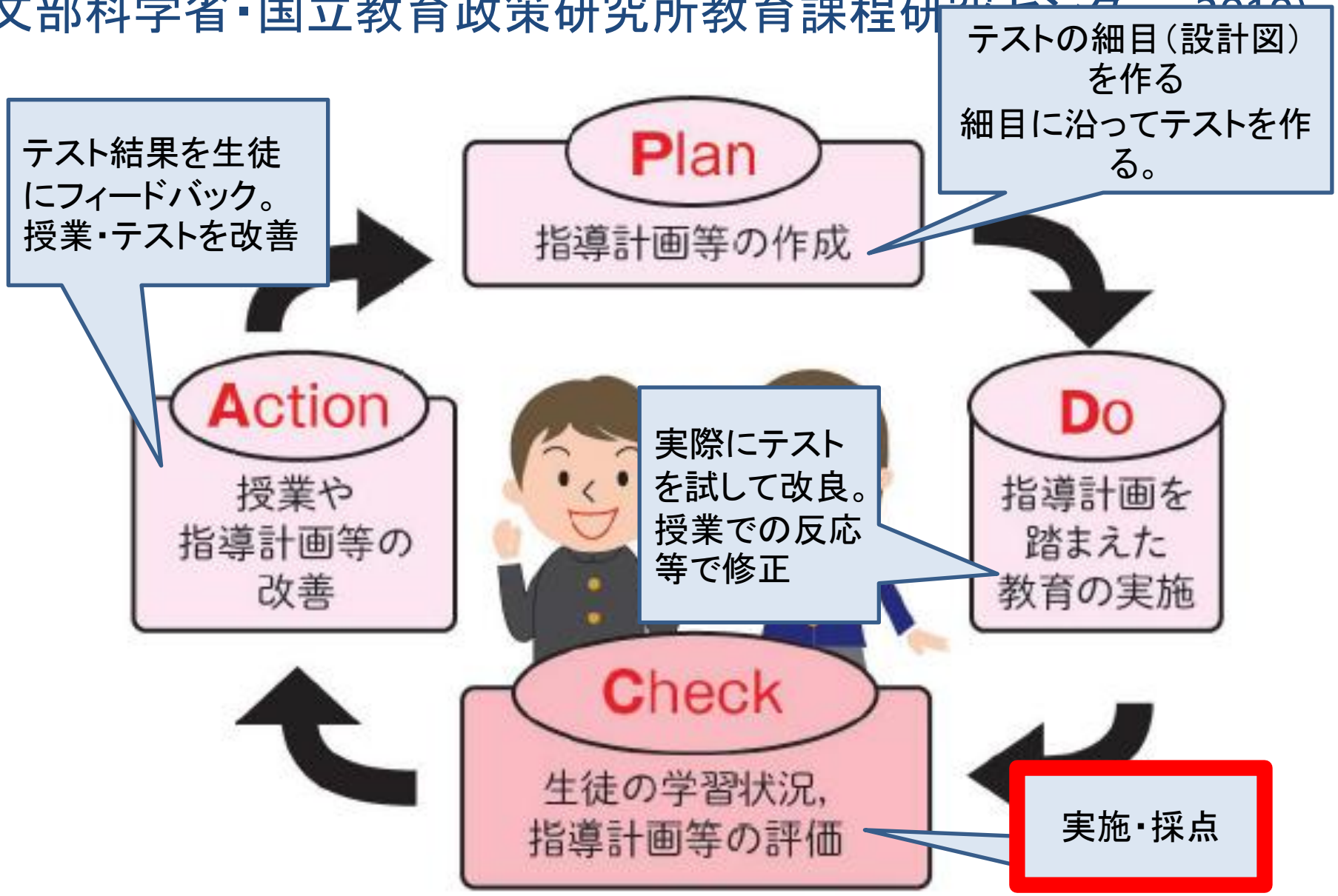
学習評価の方向性：学習のための評価

(文部科学省・国立教育政策研究所教育課程研究センター, 2019)



学習評価の方向性：学習のための評価

(文部科学省・国立教育政策研究所教育課程研究センター　2010)



生徒のS力が高まるようなテスト構成にするための6原則 (Muñoz & Álvarez, 2010に基づく)

- 1）テスト内の**タスク形式**や個々のタスクは、指導目標と指導内容と一致させる。コミュニケーション能力の多様な要素を測り、現実でも行うようなタスクを設定する
- 2）S力をうまく引き出す**指示・手順**を考える
- 3）**ルーブリック（評価基準）**は、測りたい力を明記し、教師と生徒に理解しやすいものにする。事前に示す
- 4）**自己評価**と**相互評価**を奨励する
- 5）結果の返却時には、可能な範囲で、具体的で詳細な**フィードバック**を渡す
- 6）**継続的に評価**を行う

生徒のS力が高まるようなテスト構成にするための6原則 (Muñoz & Álvarez, 2010に基づく)

-
-
-
-
- 細
- 6)**継続的に評価**を行う

採点の安定性（信頼性）は学習のための評価を支える柱

信頼性：テスト結果が十分に安定しているか

☆大規模テストと違い、授業内テストでは高い信頼性はなくてもよい。中程度は必要

例：採点者間で同じように採点しているか

例：同じ採点者が同じように採点しているか

採点者の影響を減らすための方法

- 教員間・教員内の採点の一貫性（信頼性）の確保
- 採点前の**事前確認**・トレーニング
  - 採点基準や発話例を確認した後、似た状況で行ったビデオや音声を使って、別々に評価。ずれをなくすために話し合う
  - よくある状況：長時間は取れない
- 採点中
  - （一部でも）**2名の採点者によるダブル採点**
  - よくある状況：他の教員に採点が依頼しにくい
- 採点後　**ずれを確認**。話し合ってずれの理由を特定。最終スコアを決定
  - よくある状況：あまり時間は取れない。時間があってもなぜずれたかが特定できず、修正しにくい

採点者の影響を減らすための方法

- 教員間・教員内の採点の一貫性（信頼性）の確保
- 採点前の**事前確認**・トレーニング
  - 採点基準や発話例を確認した後、似た状況で行ったビデオや音声を使って、別々に評価。ずれをなくすために話し合う
  - よくある状況：長時間は取れない
- 採点中
  - **最初の2～3名を独立に評価したところで、得点を見せ合って、基準を調整（宇都宮南高校）**
  - （一部でも）**2名の採点者によるダブル採点**
  - よくある状況：他の教員に採点が依頼しにくい
- 採点後　**ずれを確認**。話し合ってずれの理由を特定。最終スコアを決定
  - よくある状況：あまり時間は取れない。時間があってもなぜずれたかが特定できず、修正しにくい

研究課題

- Q1: シンプルなルーブリックを用い、長時間の採点者トレーニングがない場合で、採点者1名の場合に、どの程度の信頼性になるか？
- Q2: 採点者1名で低い信頼性の場合、何名いれば信頼性が確保できるか？
- 十分な信頼性：Φ = .70以上
- Koizumi & Watanabe (印刷中)
  - 事前の打ち合わせ：10分
  - 最初の2～3名を独立に評価したところで、基準を調整
  - 3観点、3段階のルーブリック
- 小泉・初澤・磯部・松岡 (2021)
  - 事前の打ち合わせ：10分
  - 3～4観点、3段階のルーブリック

分析

- Q1: シンプルなルーブリックを用い、長時間の採点者トレーニングがない場合で、採点者1名の場合に、どの程度の信頼性になるか？
- Q2: 採点者1名で低い信頼性の場合、何名いれば信頼性が確保できるか？
  - 多変量一般化可能性理論 (Grabowski & Lin, 2019; 小泉, 2018)

- 他に、採点者やルーブリックの観点も分析したが大きな問題はなし
  - 多相ラッシュ分析 (Linacre, 2021a, 2021b; 平井・横内・加藤, 2018)

研究1: Koizumi & Watanabe (印刷中)

- 高1　40名3クラス = 120名　116名分析
- 精密な採点者トレーニングはできなかった
  - 10分ほどの事前打ち合わせ。テストの最初の2・3名の採点後に話し合いのみ
- その状態でどの程度信頼性（評価者内、評価者間の安定性）が保てるか？
- 評価者
  - 第2回目:2名　小泉＋担当教員1名
  - 第3回目:3名　小泉＋担当教員1名＋同校英語教員1名ALT
  - 第4回目:9名　小泉＋担当教員1名＋同校英語教員7名
  - 第5回目:8名　小泉＋担当教員1名＋同校英語教員6名

実施形態：年6回実施

	第１回 ５月	第２回 ７月	第３回 10月	第４回 12月	第５回 １月	第６回 ２月
授業回数	2	1	2	2	1	2
テスト形式	教師面接型	個人プレゼンテーション	Group discussion	Paired oral 生徒同士の会話	Group discussion	教師面接型
測る力	やり取り	発表＋やり取り	（発表＋）やり取り	やり取り	（発表＋）やり取り	やり取り
採点	事後	実施中				
場所	教育相談室	図書館の一角				
テスト受験者以外の活動	授業同時進行	他生徒は別課題	全員同時にテスト活動	他生徒は課題	全員同時にテスト活動	他生徒は課題行う

- 成績には入れず、形成的に活用
- 結果をスコアレポートとして返却。授業で改善点等を解説。自己評価 (Koizumi et al., 2020)

スコアレポート (第1回目)

話すこと(やりとり)についてのスピーキング・テストの結果

1 年 A 組 氏名 江戸川 新一

2019 年 5 月に受けた、スピーキング・テストの結果です。これからの英語の勉強に活かしてください。

あなたの結果 (見方は一番下の表を参照)	あなたの学年の平均
A1.2	A1.3
あなたのできることの目安: よく知っていることについて、基本的な表現を使って、ごく簡単なやりとりができます。	

◎ これが得意のようです　それ以外は改善できるか考えてみましょう	あなたの特徴
1. あまりポーズなしに、文をすらすら言うことができる	
2. 文を正確に言うことができる	
3. いろいろな語句や文法を使って話すことができる	
4. 伝えたいところを強調して話すことができる	
5. 文をたくさん言ったり、話をふくらませたりしようと努力することができる	
6. 相手に聞こえるはっきりとした声で話すことができる	
7. 相づちや短い応答で積極的に聞いていることを示すことができる	
8. わからないときには "OK?" "One more time, please." など聞いたり確かめたりしながら話すことができる	
9. 笑顔、アイ・コンタクト、その場にあったジェスチャーなどができる	

CEFR-Jレベル	A1.1	A1.2	A1.3	A2.1	A2.2
英検の級の目安	5 級合格程度	4 級合格程度	3 級合格程度	準 2 級に挑戦し始める程度	準 2 級合格程度
CEFR-J の各レ	なじみのある	基本的な語や	趣味、部活動な	順序を表す表	簡単な英語で、

第1回目解説：多くが苦手だった点

- １）　聞かれた質問に答えることはできていても、**自分から質問するのは苦手**のようでした。受け身でなく、相手に質問する練習を多くするといいですね。
- ２）　好きな食べ物や趣味は答えられても、**「それについてもっと話して」と言われると困って**しまう人がいました。好きな物について、いつ、どこで、だれが、何を、なぜ、どのように、という点を加えて話を広げられるように練習するといいですね。…
- ４）　会話中に、自分の言いたいことが言えない時に、考えて黙ってしまうのでなく、**Let me seeやWell, please wait for a while.など、少し考えるよというサインを出す**方法を身につけたいですね。また相手の言ったことが聞き取れない時に、Could you say that again? など、もう一度言ってもらうように言えた人も少なかったです。多かったのは、分からなくて黙ってしまうケースでした。気をつけましょう。また、一部できていた人がいましたが、相手が言ったことについて、Oh. I see. Great. OK.など相づちを打つ練習もしたいですね。

スコアレポート (第2～5回目)

話すこと(発表)についてのスピーキング・テストの結果(スコアレポート)

1 年　番号　A　氏名　江戸川 新一

2019 年 7 月に受けた、スピーキング・テストの結果です。これからの英語の勉強に活かしてください。

(結果の見方は、表 1 を見てください)

	あなたの結果	学年の平均
タスク達成度	A	B
流暢さ	C	B
質問への応答	B	B
全体	B	B+

表 1. 準備事前した英語プレゼンテーションで、どんなことができるかの目安

	タスク達成度	流暢さ	質問への応答	全体
A	指定された 4 項目（以下参照）のうち、4 個全部の内容を十分伝えることができる。	5 秒以上の長い沈黙がない。言い直しがあっても気にならない程度である。スムーズに話すことができる。	質問に基づいた応答が満足のいく内容でできる。	プレゼンテーションとその質疑応答が**十分満足**できる形で行うことができる。
A-	Aより少しだけ足りない状態でプレゼンテーションができる			
B+	Bより少しだけ良い状態でプレゼンテーションができる			
B	2～3 個の内容を伝えることができる。	長い沈黙が 1 回ある。言い直しが多めで少し気になる程度である。遅めスピードで話すことができる。ほとんどが B の記述に当てはまる。	質問に基づいた応答ができるが、内容が不十分である。	プレゼンテーションとその質疑応答が**おおむね満足**できる形で行うことができる。
B-	Bより少しだけ足りない状態でプレゼンテーションができる			
C+	Cより少しだけ良い状態でプレゼンテーションができる			

スピーキング・テストの結果（スコアレポート）：2019 年 5 月（第 1 回）と 2020 年 2 月（第 6 回）

第6回

番号 4869 氏名 江戸川 新一

	第 1 回	第 6 回
あなたの結果 （見方は下の表を参照）	A1.3	A2.2
あなたのできることの 目安	よく知っていることについて、文を使いながら、ごく簡単なやりとりができます。	日常生活での身近なことがらについて、詳細な意見や情報のやりとりが少しできます。
あなたの学年の 平均	A1.3	A2.1

表. どんなことが英語でできるかの目安

CEFR-J	A1.1	A1.2	A1.3	A2.1	A2.2	B1.1
英検の級の目安	5 級合格程度	4 級合格程度	3 級合格程度	準 2 級に挑戦し始める程度	準 2 級合格程度	2 級に挑戦し始める程度
CEFR-J の各レベルでできること	なじみのある定型表現を使って、時間・日にち・場所について質問したり、質問に答えたりすることができる。	基本的な語や言い回しを使って日常のやりとり(何ができるかや色についてのやりとりなど)において単純に応答することができる。	趣味、部活動などのなじみのあるトピックに関して、はっきりと話されれば、簡単な質疑応答をすることができる。	順序を表す表現である first, then などのつなぎ言葉や「右に曲がって」や「まっすぐ行く」などの基本的な表現を使って，単純な道案内をすることができる。	簡単な英語で、意見や気持ちをやりとりしたり、賛成や反対などの自分の意見を伝えたり、物や人を較べたりすることができる。	身近なトピック（学校・趣味・将来の希望）について、簡単な英語を幅広く使って意見を表明し、情報を交換することができる。
	家族、日課、趣味などの個人	スポーツや食べ物などの好	基本的な語や言い回しを使	補助となる絵やものを用い	予測できる日常的な状況（郵	個人的に関心のある具体的

1. & 6. 教員面接型
2. プレゼンテーション

表 1. 準備事前した英語プレゼンテーションで、どんなことができるかの目安

	タスク達成度	流暢さ	質問への応答	
A	指定された 4 項目（以下参照）のうち、4 個全部の内容を十分伝えることができる。	5 秒以上の長い沈黙がない。言い直しがあっても気にならない程度である。スムーズに話すことができる。	質問に基づいた応答が満足のいく内容でできる。	
A-	Aより少しだけ足りない状態でプレゼンテーションができる			
B+	Bより少しだけ良い状態でプレゼンテーションができる			
B	2～3 個の内容を伝えることができる。	長い沈黙が 1 回ある。言い直しが多めで少し気になる程度である。遅めスピードで話すことができる。ほとんどが B の記述に当てはまる。	質問に基づいた応答ができるが、内容が不十分である。	
B-	Bより少しだけ足りない状態でプレゼンテーションができる			
C+	Cより少しだけ良い状態でプレゼンテーションができる			
C	1 個のみの内容を伝えることができる。	長い沈黙が 2 回以上ある。言い直しが多い。話すスピードが遅	質問に基づいた応答がほとんどできてい	

3. & 5. グループ型会話

4名でのディスカッションでのルーブリック例

（1）内容 （テーマ・話し合いの流れに沿った発言内容） 10 点満点

テーマに即した内容について、自分自身の考えや意見についてわかりやすく話している。	10
テーマに即した内容について、流れに沿った発言をして話し合いに参加しているが、内容がやや限定的、または不明瞭である。	7
テーマに即して話そうとしているが、内容がやや不明瞭である。	4

（2）やり取りの適切さ（バランスの取れた発言と他者の意見を聴く姿勢） 10 点満点

他者の意見に耳を傾けたり、既習の会話表現を積極的に用いて反応したり質問したりすることで、グループ内の会話の継続に努めている。	10
他者の意見にも耳を傾けながら、それについて自分の意見を述べようと努めている。	7
積極性には欠けるが、時々発言したり、他者の意見を聞いたりしている。	4

（3）コミュニケーションへの意欲 （準備した紙を見ずに英語で話そうとする姿勢） 5 点満点

英語のみでやりとりしようとしている。準備した紙をほとんど見ていない。	5
英語でやりとりしようと努めている。準備した紙をたまに見ている。	3
英語での発言が少ない。準備した紙を常に見ている。	1

4. ペア型会話

ペア型タスク（千菊, 2019, pp. 97-111, 5種類カード用意）

- **1 駅まで歩くか、バスに乗るか**
- 2 テントと小屋（cabin）どちらにとまるか
- 3 修学旅行はグループ単位かクラス単位か
- 4 将来子どもは、一人暮らしをするか、一緒に暮らすか
- 5 デートの行先は動物園でショー見物か、別の場所にするか

- 教科書からは離れた投げ込み教材

「正確さ」は 今回の評価には加味しませんので、文法的な誤りについては心配せずに話してください 。

	（1） 論理性	（2） 即興性	（3） 社会性
A	意見と、それを支えるのに効果的な複数の事実や理由を述べている。	相手の意見に、質問やコメントを述べ、関心を示している。	相手の意見に反論する際、相手の論拠に理解を示すなど、心情に配慮している。
B	意見と、それを支えるのに効果的な事実や理由を述べている。	相手の意見に、質問やコメントを一部述べて反応している。	相手の発言の意図を確認しながら、話を進めている。
C	意見のみを述べている。または事実や理由があっても、意	相手の意見にほとんど反応していない。	相手の発言の意図を確認しなかったり、理解を示したり

結果

・プレゼンとペア型は高い信頼性。1～2名で十分
　⇒会話がある程度予想できるタスク、別部屋で採点に集中できた場合
・グループ型は低い信頼性。3～4名必要

採点者の数と信頼性の関係

	2nd[a] プレゼン型 採点者 2 名必要				3rd[b] グループ型 採点者 4 名必要				4th[b] ペア型 採点者 1 名でよい				5th[b] グループ型 採点者 3 名必要			
R	C1	C2	C3	計	C1	C2	C3	計	C1	C2	C3	計	C1	C2	C3	計
1	.69	.45	.71	.69	.17	.40	.35	.39	.79	.80	.82	.86	.10	.34	.54	.45
2	.82	.63	.83	.82	.29	.57	.51	.56	.88	.89	.90	.93	.18	.50	.70	.62
3	.87	.71	.88	.87	.38	.67	.61	.66	.92	.92	.93	.95	.25	.60	.78	.71
4	.90	.77	.91	.90	.45	.73	.68	.72	.94	.94	.95	.96	.31	.67	.83	.77
5	.92	.81	.92	.92	.50	.77	.73	.76	.95	.95	.96	.97	.36	.72	.86	.80

Note. R = Rater. C1 = The first criterion. To. = Total. Underlined = Φ = .70 or more.
[a]Using the data of Raters 1 and 2. [b]Using the data of Rater 2 and Rating 3.

グループ型採点が安定しない理由

・1) 各採点者は異なるパフォーマンスを採点している（トピック、順序の影響）
・2) 認知的負荷が高い（誰がいつ話すか、また発話内容が予想できない中で、1グループ3～4分で4人を同時採点）
・3) 騒音、声の大きさの違い
・4) 最初の議論を、異なるパフォーマンスの採点後に行う
・5) 測る力の把握が難しい

- C1. 内容　C2. やり取りの適切さ　C3. コミュニケーションへの意欲
- C1. 会話全部を聞く必要があり、難しい
- C2. 話し合いがあまりないと、やり取り力にばらつきが出てくる
- C3. 3つ目の観点は抜けがち

研究2：小泉・初澤・磯部・松岡(2021)

- 40名6クラス 225名分析
- 精密な採点者トレーニングはできなかった
  - 10分ほどの事前打ち合わせのみ
- その状態でどの程度信頼性（評価者内、評価者間の安定性）が保てるか？
- テスト2回分　Group discussionとGroup debate
- 評価者
  - 生徒1名を2名の教員で採点
  - 第1～3回目：担当教員1名＋同校英語教員2名

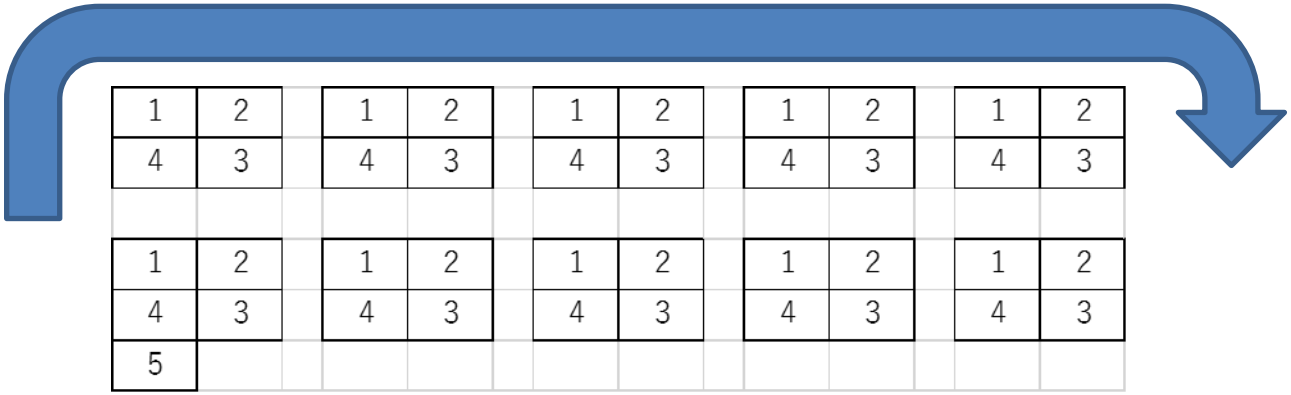
実施形態

3 学年 2 科目でのテスト形式一覧

	第 1 回 6 月	第 2 回 9 月	第 3 回 11 月	第 4 回 2 月
高 1：コミュニケーション英語 I	生徒 2 名での pair talk			
高 1：英語表現 I	個人で Show & Tell 形式			
高 2：コミュニケーション英語 II	3 人 1 グループでの chain talk	左に同じ	同じ	4 人 1 グループでの chain talk
高 2：英語表現 II	教員と生徒 1 名の面接形式	個人で Show & Tell 形式	3 人 1 グループで面接形式	3 人 1 グループでスピーチ形式
高 3：コミュニケーション英語 III	3 人グループで議論（注 1）	個人面接型	4 人グループでディベート	
高 3：英語表現 III	個人プレゼンテーション（注 1）	4 人で 5 分程度のディスカッション形式	3～4 人でのグループワーク（アンケート結果発表）	

第2回9月ディスカッション

- ４人で5分程度。各グループが10回行う。
- Moderator(司会者)は、１回のディスカッション毎に交代し、合計２～３回担当することになる。
- Moderator以外はParticipants(参加者)として意見を述べるが、Moderatorも意見を述べても構わない。
- ★Participantsの発言：１人目→２人目→３人目→(４人目)→１人目の順とする。参加者の１人目も直前の参加者の意見に言及する。



Performance Test　Evaluation Sheet　English Expression Ⅱ

[Moderator]

Class:　　No.　　Name

内容	まとめ方	参加者の発言を別の表現で言い換え、要約できている。
		参加者の発言を繰り返すなどして、適宜フィードバックしている。
		参加者の発言に言及しようとしているが、不十分である。
表現	発音・音声・ノンバーバルコミュニケー	適切にアイコンタクトやジェスチャーなどを用いており、自然に議論を進行することができてい
		時折アイコンタクトやジェスチャーを用いており、自然な議論の進行に努めている。
		アイコンタクトやジェスチャーが不十分であり、議論の進行に滞りがある。
技術点	文法・語法	文法・語法のエラーが少なく、既習の表現を活用しながら自分の考えを不足なく伝えられている
		文法・語法について多少のエラーや単純な表現は多いが、自分の考えが伝えられている。
		文法・語法のエラーが目立ち、自分の考えが十分に伝えられていない。

[Participant 2]

Class:　　No.　　Name

内容	考え方	相手の考えと自分の考えを比較しながら、根拠に基づいた自分の考えを述べている。
		根拠に基づいた自分の考えをわかりやすく述べている。
		自分の考えを述べているが、内容がやや不足である。
表現	発音・音声・ノンバーバル	適切にアイコンタクトやジェスチャーなどを用いており、自然な（理解されやすい）表現をして
		時折アイコンタクトやジェスチャーを用いており、自然な（理解されやすい）表現に努めている

第3回11月ディスカッション

- (1) 授業担当者が事前に4人のグループを作成（1クラス約10グループ）
- (2) テスト当日、授業担当者が2つのグループとトピック、また肯定側・否定側をくじで決める。
- (3) ２グループの試合形式で、それぞれのチームが以下の順序で議論を進める。（オーディエンスの前で実施）

第3回11月ディスカッション

	Stage	Team A	Team B
準備４分	① 肯定側立論(90sec)	論題を肯定する立場でメリットを述べる。	
	② 否定側立論(90sec)		論題を肯定する立場でメリットを述べる。
準備2分	③ 質疑(60sec)	相手側チームの意見に質問をする。	
	④ 質疑(60sec)		相手側チームの意見に質問をする。
準備2分	⑤ 反論(90sec)	立論・質疑を踏まえた上で、再度自分たちの優位性を説明する。	
	⑥ 反論(90sec)		立論・質疑を踏まえた上で、再度自分たちの優位性を説明する。
準備2分	⑦ 総括(90sec)	自分たちの主張の方が重要性が大きいことを印象付ける。	
	⑧ 総括(90sec)		自分たちの主張の方が重要性が大きいことを印象付ける。

グループ名

Performance Test　Evaluation Sheet　English Expression Ⅱ

【肯定側立論　／　否定側立論】(Affirmative Constructive Speech　／ N

Class:　　No.　　Name

内容	12	トピックについて、自分の立場に即して考えや意見を理由や具体例を添えて
	8	トピックについて、自分の立場に即して考えや意見を説明している。
	4	トピックについて、自分の立場に即して考えや意見を説明しようとしているが、
コミュ姿勢	8	自然なアイコンタクトやジェスチャーを用いている。積極的に発言して議論の
	5	アイコンタクトやジェスチャーを用いる努力がある。時々言いよどみながらも議
	3	アイコンタクトやジェスチャーが不十分である。積極性に欠け、議論への参加
文法語法	5	理解を妨げるような文法・語法のエラーがほぼなく、自分の考えを不足なく伝
	3	理解を妨げるような文法・語法のエラーが少しあり、自分の考えが十分に伝
	1	理解を妨げるような文法・語法のエラーが多く、自分の考えが十分に伝えら
音量速度発音	5	話す音量や速さを調整し、また発音にも留意しながら、聞き手が理解しやす
	3	話す音量や速さを調整し、聞き手が理解しやすいように努めている。
	1	話す音量や速さが、聞き手にとって理解しずらいものである。

12点　We think that Japanese high schools should allow their students to work part-t school fees, some to buy essential goods for their school life, and others to sup

8点　We think that Japanese high schools should allow their students to work part-t through working part-time. The experience of part-time job will be helpful for t

【質疑】(Question)

Class:　　No.　　Name

内容	12	相手側の主張を的確に要約した上で、相手側の立論で提起した内容につい
	8	相手側の立論で提起した内容について質問している。
	4	相手側の立論の内容について質問しようとしているが、内容が不明瞭である
	8	自然なアイコンタクトやジェスチャーを用いている。積極的に発言して議論の

結果

・両方とも高い信頼性。1名で十分
　⇒グループ型でも、話す順番等が決まっているなどで会話がある程度予想できるタスク
　⇒研究1との相違点：採点者全員が同じ学年、同教材で指導・評価。日ごろからの教材の共有

採点者の数と信頼性の関係

	第 2 回グループ型				第 3 回グループ型				
	採点者 1 名必要				採点者 1 名必要				
R	C1	C2	C3	計	C1	C2	C3	C4	計
1	.62	.67	.59	.81	.74	.68	.66	.56	.79
2	.77	.80	.74	.89	.85	.81	.79	.72	.88
3	.83	.86	.81	.93	.90	.86	.85	.79	.92

Note. R = Rater. C1 = The first criterion. To. = Total.
Underlined = Φ = .70 or more.

高校の教室内スピーキングテストにおける採点

- シンプルなルーブリックを使えば、長時間の評価者トレーニングは必要ない場合もある
  - 会話の流れや内容が予測可能な場合
  - 日ごろから教材やタスクを共有し、どの程度の発話を求めるかの共通認識がある場合
- それでも、ある程度事前、テストの最初のすり合わせができるとよい
- テストの最初に行う利点：前年の生徒のパフォーマンスを使って話し合っても、今回タスクやルーブリックを修正したり、生徒の反応が異なる場合もある
- 別な場合での研究分析をしてさらに指針作成に活かす

プロジェクト報告2：高校の英語授業内スピーキング評価における「信頼性確保のための採点指針」の作成に向けて

順天堂

JLTA

小泉利恵科研プロジェクト第2回例会

(共催：日本言語テスト学会)

2021年2月14日（日）　15:10～15:40

小泉利恵（順天堂大学）・深澤真（琉球大学）・横内裕一郎（弘前大学）・井上千尋（University of Bedfordshire）

科研プロジェクト

- 1. 国内外のスピーキング評価の理論や実践のまとめ（特にルーブリックについて）
- 2. 日本での教室内スピーキング評価の課題に関する研究
  - 報告1
  - スコアレポートのフィードバックの研究 (Koizumi et al., 2020)
  - 注意点などをまとめた指針の作成
- 3. スピーキング評価の情報をまとめたウェブサイトの公開
  - 採点トレーニングに使えるような情報を充実させる
- 4. ウェブサイトの有効性について研究

3. スピーキング評価の情報をまとめたウェブサイトの公開

- 目的
- 日本で英語を教える教員が参照できるウェブサイトを作る
- 使用例
  - トレーニング用、参照教材として
  - 個人で
  - 学校での英語科会議で
  - 地域での研修で

ニーズ分析

- 文科省の調査から
- スピーキングテストの実施頻度は、高校は多くない
- 使用タスクに偏りがある
  - やり取りの生徒会話型が少ない

- 髙島 (2019)：中学校の教員から：発表よりやり取りの評価に自信がない (自信がない人：やり取りで57.1%)
- 金子 (2019)：教員からの声「参照できる資料が欲しい」「中学校では、ALTにまかせている場合が多い」

ニーズ分析

- 2つのセミナー参加者対象アンケートから (n = 24, 30)

ニーズ分析から：課題

- コロナ禍前は、テストの採点
- コロナ禍では、テストの実施
- スピーキング評価の各段階で課題を感じている方あり

テストの作成

テストの実施

テストの採点

結果のフィードバック

結果に基づき指導の改善

単発でなく、定期的に実施すること

同僚との意見のすりあわせ（例：学年で統一して実施すること）

保護者の理解を得ること

スピーキング評価の情報をまとめたウェブサイト

- 言語：日英
- 構成
- 1) スピーキング評価全般
  - 指導用タスクは含めない
- 2) オンライン・トレーニング・サイト
  - タスクとルーブリック、サンプルパフォーマンス（極力ビデオ）
  - やり取りタスク、発表タスク
  - 準備度が高いタスク、即興性が高いタスク
  - 独立型、技能統合型

モデル

- Tools to Enhance Assessment Literacy (TEAL). (2020). TEAL for teachers of English as an additional language. https://teal.global2.vic.edu.au/

Assessment Tools

Common Writing Assessment Tool

- Guiding principles
- TEAL writing assessment tasks and criteria

Common Oral Assessment Tool

- Guiding principles
- TEAL oral assessment tasks and criteria

Common Reading & Vocabulary Assessment Tool

- Guiding principles
- The RVEAL assessment tool is now available!

Other Useful Assessment Tools

- The Reading self-assessment tool for students
- TEAL writing task observation guide

Guiding principles and advice

1. Preparing to assess students' learning
2. Getting students ready for an assessment
3. Choosing an oral assessment task
4. Using the assessment criteria
5. Making trustworthy assessment decisions
6. Aligning your assessment with the EAL Continuum
7. Using the assessment data to improve learning
8. Glossary of terms used in the TEAL criteria sheets

Oral assessment tasks and criteria

Please click on the links in the matrix below to view information on each common oral assessment task.

The TEAL text framework explains the nature of informative, imaginative and persuasive texts.

Task Name	Informative	Imaginative	Persuasive
Task 1: My weekend	A1, A2, BL, B1, B2, SL, S1, S2		
Task 2: What did you have for lunch?	A1, A2, BL, B1, B2, SL, S1, S2		
Task 3: Story retell		A1, A2, BL, B1, B2, SL, S1, S2	
Task 4: How to make an origami dog			A1, A2, BL, B1, B2, SL, S1, S2
Task 5: Building a bridge			A1, A2, BL, B1, B2, SL, S1, S2
Task 6: What I learn at school	A1, A2, B1, B2, S1, S2		
Task 7: Playing a card game: Fish			A1, A2, BL, B1, B2, SL, S1, S2
Task 8: Picture sequence narrative		A1, A2, BL, B1, B2, SL,	

Task 13: Choosing a gift for a character		B2, B3, S2, S3, S4	
Task 14: Description of a cycle	A2, B1, B2, B3, S1 , S2, S3		
Task 15: Survival items			S2, S3, S4
Task 16: Note taking from a telephone conversation	S2, S3, S4		
Task 17: If I won a million dollars		B2, B3, S2, S3, S4	
Task 18: Role play: Giving advice to a friend			B2, B3, S2, S3, S4
Task 19: A book or film review	B3, S2, S3, S4		
Task 20: A sales talk			B3, S2, S3, S4
Task 21: Job interview role play			B3, S2, S3, S4

18. Role play: giving advice to a friend

- Task details
- Task specification
- Learning/teaching context
- Task implementation
- Assessment criteria
- Samples of student work
- Annotations and commentary

Task procedure: **(Instructions for students)** and stimulus/prompt: **(rubric,visuals)**

1. Student ‘A’ and ‘B’ are friends
2. Student ‘A’ begins by explaining the problem and asking for advice
3. Student ‘B’ listens to the problem and responds talking about his / her own experience
4. Students ‘A’ and ‘B’ discuss the situation and ideas about what to do next
5. You have 5 to 7 minutes to complete the role-play.

TEAL Oral assessment criteria **Task 18: Giving advice to a friend** **Name:_____** **Date: _____**

	Communication	Cultural conventions of language use	Linguistic Structures and features and intelligibility				Strategies	AusVELS EAL Stages		
			Text structure	*Grammatical features*	*Vocabulary*	*Phonology*			B2	B3
								S2	S3	S4
4	• Describes problem (student seeking advice) and suggests solution (student giving advice) • Explores options (both students) • Suggests choices (student giving advice) • Discusses of criteria for choosing between options (student giving advice)	• Relaxed interactional style • Empathetic exploration • Facial expressions	• Greeting • Request for help • Statement of the problem • Recount of previous experience • Suggested optional solutions • Exploration of problems and suggestions • Response to suggestions	• Uses a range of structures to ask questions, describe and explore suggestions, ask questions, comment on • Uses a range of tenses to describe the problem, past events and a range of structures for suggesting *I suggest…You could, why.. Why not..?* • Uses a range of structures to talk about definite future events – present tense, *going to* and *will* Conditional if… • Use of range of modals for making suggestions, including question form – *Should I pick this*	• Uses a range of adverbs to structure comments – *overall*, add emphasis – *I'm really excited, That's a really tough question,* or qualify *Then just go..* • Uses a range of words for the topic of conversation • Uses modal adverbs – *perhaps, maybe*	• Highly intelligible, • Clear articulation of sounds • Varied intonation coveys interest and attitudes of speakers, including occasional frustration	• Makes extensive use of hand gestures and facial expressions to add emphasis and support meaning			
3	• Describes problem and makes explicit appeal for help (student seeking advice) • Suggests solutions (student giving advice) • Explores implications of suggestions (student giving advice) • Comments on aspects of the situation or solutions (student giving advice)	• High degree of Interaction, often short interactive turns and comment or response • Nodding acknowledgement of partner, back-channelling	• Greeting • Opening request for help • Statement of the problem • Recount of previous experience • Suggested solution • Exploration of ideas • Response to suggestions	• Uses more complex and complete sentence structures, although not always fully accurate • Uses longer and more complex verb phrases *you should think about…? Have you thought about…?* • Uses simple past for previous events, not consistent • Uses a range of structures for definite future events – present tense, *going to* and *will* • Uses some modals to make suggestions – *you should, you can*	• Uses words related to school – *home room teacher,* • Uses words for family, relatives • Uses adverbs *actually, really, pretty (sure)* • Makes some use of modal adverbs – *maybe*	• Intelligible pronunciation • Some missing elements – final consonants • Varied intonation to show empathy and interest	• Uses gestures, head movements to add emphasis and interest to role play			
2	• Describes and states problem (student seeking advice) • Suggests solutions (student giving advice • Explores consequences of some actions (student giving advice) • Makes final suggestion (student giving advice) and signals acceptance of solution	• At ease in interaction, responsive turn-taking • Responsive, fluent, empathetic interaction • Nodding, short interjections – to signal understanding	• Statement of problem • Questions to understand the situation more and explore options • Discussion of possible consequences • Suggested solution	• Uses more complete structures – *I think you really want to* • Uses well-formed questions *Have you thought about getting your relative to come to the concert?* • Uses simple past for past events • Uses simple present and *will* for marking likely future events • Uses modals for giving advice, *you could, you should* as well as *can*	• Uses *because* to give reasons • Uses *How about if* for speculating • Uses conjunction *but* to relate ideas	• Some elements missing – final consonants • Some sounds not articulated accurately – ***parents*** sounds like ***barents*** • Intelligible, influence of L1 is noticeable	• Uses some helping phrases to keep the conversation going			
1	• Describes the problem and requests help (student seeking advice) • Suggests single solution with reasons (student giving advice) • Explores some reasons and consequences of the suggestion (student giving advice)	• Turn taking is rather formal, indicated by pauses rather than by 'hand over' strategies • Distinct turns with limited interaction • Limited expression of empathy until completion of the task	• Statement of problem • Statement of suggested solution • Expression of thanks	• Uses limited or not always complete sentence structures – *I don't want* • Uses some incomplete verb phrases - *he will only angry* • Uses simple present for past events • Uses simple present for likely future events • Uses future to indicate hypothetical meaning expressed – *you can tell* • Uses modal *can* for suggestions • Makes some structural errors e.g. questions *How can I do?*	• Uses vocabulary for the situation – *school concert* • Uses comparatives – *.more than, more important than* • Makes some preposition errors – *same time* ***with****, angry* ***to*** *me,* by a long time	• Not all sounds clearly articulated, and some words hard to identify – *relative more than,* • Limited linking, sounds slightly staccato	• Uses some gestures, e.g. nodding to add emphasis • Limits discussion, does not refer to previous experience • Defers to partner's preference and priority to end discussion			

Sample One | Sample Two | Sample Three | Sample Four | Samp

Sample 1

Biographical Information

Year 10 (both students)

Student A (left of screen) Thai

Student B (right of screen) Korean

Commentary

These students are able to communicate on the topic, despite limitations and noticeab
control over some features, such as the modals *could* and *should* for giving advice, tha
state the problem, but uses *my parents say that* [the family dinner] *is good and impo*
want.., and asks for help. However while her meaning is clear she has problems formi
Student B is able to offer advice using better-structured utterances, *because you prac*
limited in what she says in justification for her advice, *I just follow your opinion beca*

Student A makes errors of sentence formation and the use of prepositions, as well as u
problem and reacting to her partner's suggestion. These are noticeable but she is still

Both students' pronunciation is intelligible but not always clear. Student A says some
relative and *because my school concert is my* (sounds like) '*skill*', and while student B
slightly staccato dues to errors of rhythm and linking of sounds. They are rather over p
they relax at the very end of the task. This is probably the result of the effort they are p
need to communicate their meanings, preventing them from them being more relaxed

TEAL Oral assessment criteria **Task 18: Giving advice to a friend** **Name: Sample 1** **Date: _____**

	Communication	Cultural conventions of language use	Linguistic Structures and features and intelligibility				Strategies	AusVELS EAL Stages		
			Text structure	*Grammatical features*	*Vocabulary*	*Phonology*		B2 S2	B3 S3	S4
4	• Describes problem (student seeking advice) and suggests solution (student giving advice) • Explores options (both students) • Suggests choices (student giving advice) • Discusses of criteria for choosing between options (student giving advice)	• Relaxed interactional style • Empathetic exploration • Facial expressions	• Greeting • Request for help • Statement of the problem • Recount of previous experience • Suggested optional solutions • Exploration of problems and suggestions • Response to suggestions	• Uses a range of structures to ask questions, describe and explore suggestions, ask questions, comment on • Uses a range of tenses to describe the problem, past events and a range of structures for suggesting *I suggest…You could, why.. Why not..?* • Uses a range of structures to talk about definite future events – present tense, *going to* and *will* Conditional if… • Use of range of modals for making suggestions, including question form – *Should I pick this*	• Uses a range of adverbs to structure comments – *overall*, add emphasis – *I'm really excited, That's a really tough question,* or qualify *Then just go..* • Uses a range of words for the topic of conversation • Uses modal adverbs – *perhaps, maybe*	• Highly intelligible, • Clear articulation of sounds • Varied intonation coveys interest and attitudes of speakers, including occasional frustration	• Makes extensive use of hand gestures and facial expressions to add emphasis and support meaning			
3	• Describes problem and makes explicit appeal for help (student seeking advice) • Suggests solutions (student giving advice) • Explores implications of suggestions (student giving advice) • Comments on aspects of the situation or solutions (student giving advice)	• High degree of Interaction, often short interactive turns and comment or response • Nodding acknowledgement of partner, back-channelling	• Greeting • Opening request for help • Statement of the problem • Recount of previous experience • Suggested solution • Exploration of ideas • Response to suggestions	• Uses more complex and complete sentence structures, although not always fully accurate • Uses longer and more complex verb phrases *you should think about…? Have you thought about…?* • Uses simple past for previous events, not consistent • Uses a range of structures for definite future events – present tense, *going to* and *will* • Uses some modals to make suggestions – *you should, you can*	• Uses words related to school – *home room teacher,* • Uses words for family, relatives • Uses adverbs *actually, really, pretty (sure)* • Makes some use of modal adverbs – *maybe*	• Intelligible pronunciation • Some missing elements – final consonants • Varied intonation to show empathy and interest	• Uses gestures, head movements to add emphasis and interest to role play			
2	• Describes and states problem (student seeking advice) • Suggests solutions (student giving advice • Explores consequences of some actions (student giving advice) • Makes final suggestion (student giving advice) and signals acceptance of solution	• At ease in interaction, responsive turn-taking • Responsive, fluent, empathetic interaction • Nodding, short interjections – to signal understanding	• Statement of problem • Questions to understand the situation more and explore options • Discussion of possible consequences • Suggested solution	• Uses more complete structures – *I think you really want to* • Uses well-formed questions *Have you thought about getting your relative to come to the concert?* • Uses simple past for past events • Uses simple present and *will* for marking likely future events • Uses modals for giving advice, *you could, you should* as well as *can*	• Uses *because* to give reasons • Uses *How about if* for speculating • Uses conjunction *but* to relate ideas	• Some elements missing – final consonants • Some sounds not articulated accurately – *parents* sounds like ***barents*** • Intelligible, influence of L1 is noticeable	• Uses some helping phrases to keep the conversation going			
1	• Describes the problem and requests help (student seeking advice) • Suggests single solution with reasons (student giving advice) • Explores some reasons and consequences of the suggestion (student giving advice)	• Turn taking is rather formal, indicated by pauses rather than by 'hand over' strategies • Distinct turns with limited interaction • Limited expression of empathy until completion of the task	• Statement of problem • Statement of suggested solution • Expression of thanks	• Uses limited or not always complete sentence structures – *I don't want* • Uses some incomplete verb phrases - *he will only angry* • Uses simple present for past events • Uses simple present for likely future events • Uses future to indicate hypothetical meaning expressed – *you can tell* • Uses modal *can* for suggestions • Makes some structural errors e.g. questions *How can I do?*	• Uses vocabulary for the situation – *school concert* • Uses comparatives – *.more than, more important than* • Makes some preposition errors – *same time **with**, angry **to** me,* by a long time	• Not all sounds clearly articulated, and some words hard to identify – *relative more than,* • Limited linking, sounds slightly staccato	• Uses some gestures, e.g. nodding to add emphasis • Limits discussion, does not refer to previous experience • Defers to partner's preference and priority to end discussion			

Selected criteria = Student A (left of screen) Selected criteria = Student B (right of screen) Selected criteria = Both students

ウェブサイト作成に向けて

- TEALとの違い
  - スピーキング評価のみ
  - 日本の小中高で使えるタスク・ルーブリックを意識
  - タスクの記述等は簡素化
    - 詳細を載せるときには、クリックすると読める形にする
  - ルーブリックは、TEALのように観点が決まっていないため、いろいろなあり得て適切なものを掲載
  - 日本人英語学習者のサンプル、発話書き起こしを載せる
- 未定：　どんなカテゴリー分けにするか
  - やり取りタスク（教師面接型、ペア型会話）、発表タスク（プレゼン、絵描写）などタスク別に
  - 小中高に分けて

参考文献

- Eckes, T. (2019). Many-facet Rasch measurement: Implications for rater-mediated language assessment. In V. Aryadoust & M. Raquel (Eds.), Quantitative data analysis for language assessment (Vol. I: Fundamental techniques; pp. 153–175). New York, NY: Routledge.
- Grabowski, K. C., & Lin, R. (2019). Multivariate generalizability theory in language assessment. In V. Aryadoust & M. Raquel (Eds.), Quantitative data analysis for language assessment (Vol. I: Fundamental techniques; pp. 54–80). New York, NY: Routledge.
- 平井明代・横内裕一郎・加藤剛史 (2018). 「項目応答理論」 平井明代 (編著). 『教育・心理・言語系研究のためのデータ分析—研究の幅を広げる統計手法』(pp. 94–137). 東京図書
- 金子淳 (2019). 『中高連携を踏まえた、英語授業におけるアクティヴィティとパフォーマンス・テスト開発に関する調査研究』(公益財団法人やまがた教育振興財団 平成30年度「教員養成に関する調査研究事業」調査研究報告書). https://www.gakushubunka.jp/scholarship/kenkyugaiyou_kaneko_h30.pdf

参考文献

- 小泉利恵 (2018). 「一般化可能性理論」 平井明代 (編著). 『教育・心理・言語系研究のためのデータ分析—研究の幅を広げる統計手法』 (pp. 65–93). 東京図書
- 小泉利恵・初澤晋・磯部礼奈・松岡京一 (2021). 「日本の高校における教室内グループ・ディスカッションとディベートにおける採点者信頼性」 未発表原稿.
- 小泉利恵・印南洋・深澤真 (編著) (2017). 『実例でわかる英語テスト作成ガイド』 大修館書店
- Koizumi, R., & Watanabe, A. (in press). Rater reliability in classroom speaking assessment in a Japanese senior high school. Annual Review of English Language Education in Japan, 31.
- Koizumi, R., Watanabe, A., Fukazawa, M., & Inoue, C. (2020, November). Examining learner perception toward test feedback in classroom-based speaking assessment using a validity framework. Paper presented at the LTRC/ALTAANZ Online Celebratory Event 2020.

参考文献

- 国立教育政策研究所 (2020). 『指導と評価の一体化」のための学習評価に関する参考資料（中学校　外国語）』 https://www.nier.go.jp/kaihatsu/shidousiryou.html
- Linacre, J. M. (2021a). A user’s guide to FACETS Rasch-model computer programs: Program manual 3.83.5. http://www.winsteps.com/manuals.htm
- Linacre, J. M. (2021b). FACETS: Many-Facet Rasch-measurement (Version 3.83.5) [Computer software]. Chicago: MESA Press.
- 文部科学省・国立教育政策研究所教育課程研究センター (2019). 「学習評価の在り方ハンドブック（高等学校編）」 https://www.nier.go.jp/kaihatsu/shidousiryou.html
- Munóz, A. P., & Álvarez, M. E. (2010). Washback of an oral assessment system in the EFL classroom. Language Testing, 27, 33 – 49. https://journals.sagepub.com/doi/10.1177/0265532209347148 (Open access)

参考文献

- 文部科学省 (2019). 「平成30年度「英語教育実施状況調査」の結果について」 https://www.mext.go.jp/a_menu/kokusai/gaikokugo/1418086_00001.htm
- 文部科学省 (2020). 「令和元年度「英語教育実施状況調査」の結果について」 https://www.mext.go.jp/a_menu/kokusai/gaikokugo/1415042.htm
- 千菊基司 (2019). 『即興的に話す交渉力を高める！中学校英語スピーキング活動アイデア＆ワーク』明治図書
- 髙島健治 (2019). 『思考力を働かせる学習サイクルの構築―定型表現で終わらない［やり取り］の力を育成するために―』平成30年度千葉県長期研修研究報告書
- Tools to Enhance Assessment Literacy (TEAL). (2020). TEAL for teachers of English as an additional language. https://teal.global2.vic.edu.au/