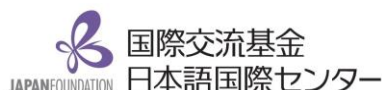


## 学習を評価する

### Unit 1 準備編 Part 2 「良いテスト」とは



#### はじめに

テストは多くの学校で学習者の日本語力を測るために行われ、とても重視されています。このパートではテストの役割、テストで何を測るのかについて確認し、良いテストを作るための観点を理解しましょう。

#### ◇キーワード

学習目標、<sup>だとうせい</sup>妥当性、<sup>しんらいせい</sup>信頼性、<sup>しんせいせい</sup>真正性、<sup>はきゅうこうか</sup>波及効果、<sup>じつようせい</sup>実用性

#### 1. 到達度テストを行うために必要なこと

到達度テストとはコースの中で行われるものですが、何のために行うのでしょうか。図1を見てください。テストの結果を見て学習者に成績をつける資料にする(①)という人が多いかもしれませんが、学習者がどの程度学習を進めているのかという学習状況を知ること(②)や、指導の成果と課題を知って(③)教授活動に役立てるといった目的もあるでしょう。

テストがこれらの目的を果たすためには、「テストの結果を分析・解釈すること(図1のA)」

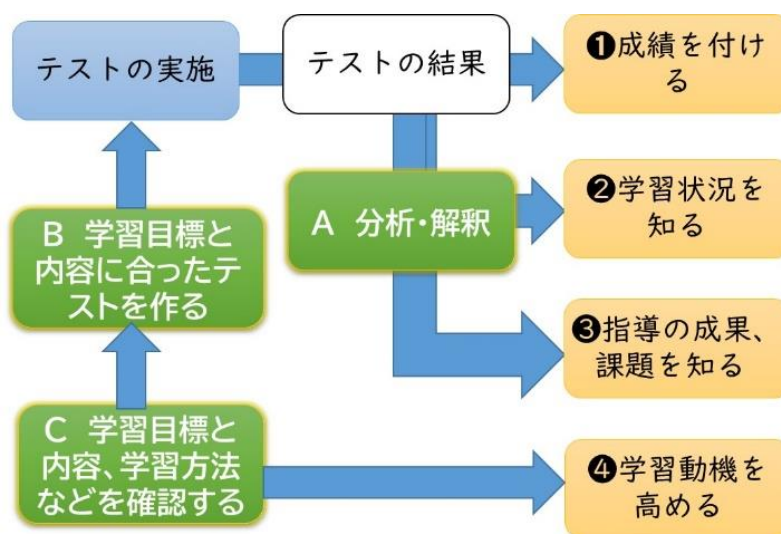


図1 到達度テストを行うために必要なこと

が必要ですが、一番大切なことは、テストが「コースの目標と内容に合っていること(B)」です。ですから、テストを作る時には、テストが「コースの学習目標と学習内容、それに合った学習の方法を確認する(C)」ものだということを忘れてはいけません。これらの点は、コースを始める前に教師が考えておくことですが、テストを準備している時にもよく確認しましょう。

また、テストによって学習者の「④学習動機を高めること」も期待できます。テストのための学習が自発的で、より良い学習になるためにも、教師はコースの初めに目標、内容、方法を学習者にしっかり伝えておく必要があります。

【タスク1】 図1の A・B・C のうち、今まで自分がテストを作る時に行っていたことに○、行っていなかったことに×を付けてください。

## 2. 学習目標を確認する

到達度テストを作る時にまず必要なことは、学習目標を確認することです。JF日本語教育スタンダード(JFS)の木を使って見てみましょう。JFSの木(図2)はコミュニケーションに必要な力を整理したものです。

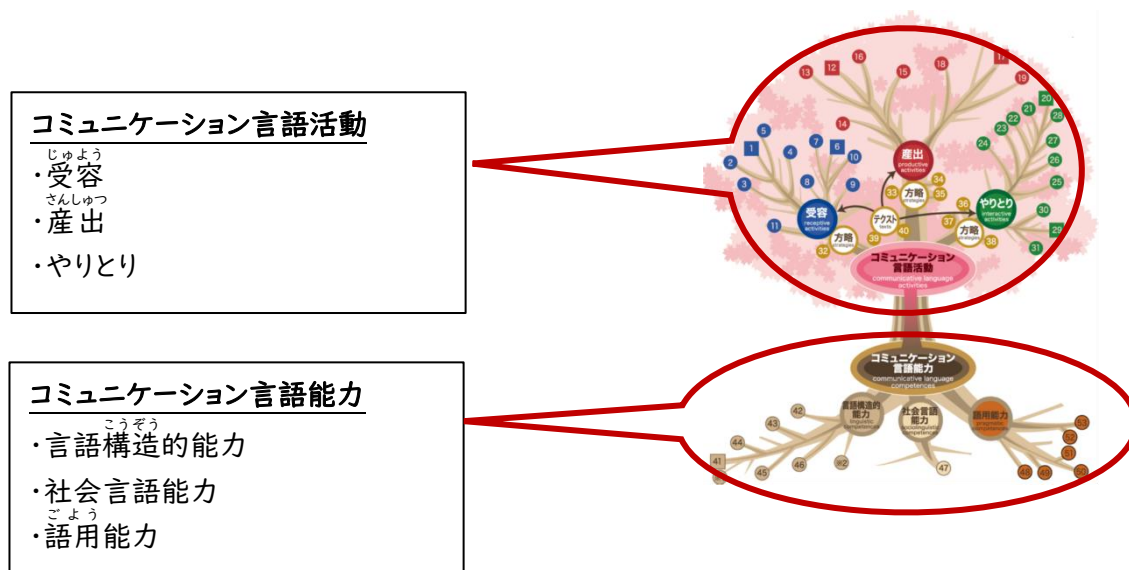


図2 JFSの木

JFSの木の枝<sup>えだ</sup>には、目に見える「コミュニケーション言語活動」の категорияが表されています。コミュニケーション言語活動は、受容(話を聞いたり、文章を読んだりする)、やりとり(相手との会話や手紙など)、産出(一人で長く話したり書いたりする)の3つに大きく分けられます。一方、JFSの木の根<sup>ね</sup>は、言語活動を支える「コミュニケーション言語能力」として表されています。言語構造的な能力・社会言語能力・語用能力など、目に見えない知識<sup>ちしき</sup>がコミュニケーションを支える力と考えます。

このJFSの木のどの部分が学習目標なのか確認することが必要です。ただし評価の対象

になるものがJFSの木にすべて表されているわけではありません。JFS の木に表現されていないけれども評価の対象となる大切な力については、Unit2のPart3で取り上げます。

【タスク 2】 あなたが担当している授業やコースの学習目標は JFS の木 (図2) の枝 (コミュニケーション言語活動) と、根 (コミュニケーション言語能力) のどちらにあたるでしょうか。考えてみましょう。

### 3. 良いテストを作るための観点

測りたい力を的確<sup>てきかく</sup>に測ることができるテストを作るための観点を確認します。

#### 1) 妥当性<sup>だとう</sup>

学習目標と内容、そして学習方法を確認したら、それを測ることができるテストを作らなければなりません。この「テストが測るべき力を測っているか」を妥当性と言います。例題を見ながら妥当性について考えましょう。

例題1は、形容詞の意味が分かっているかどうかを測るために作られたものです。

例題1 正しいものを選んで○をつけなさい。

- 1) 大阪は(たいへんな・むずかしい・にぎやかな)町です。
- 2) 富士山は(ひくい・しろいな・ゆうめいな)山です。
- 3) 日本語の勉強は(たいへん・にぎやか・べんり)です。

正しい形容詞を選ぶ問題ですが、この問題で形容詞の意味が分かっているかどうかを正しく測ることはできるでしょうか。形容詞の意味が理解できていても、例題1の1)で、大阪がにぎやかな街なのかどうか、2)では富士山が低いのかどうかを知らない学習者がいたら、正しく答えられないおそれがあります。これでは「測るべきことを測っている妥当性のあるテスト」とは言えません。妥当性のあるテストを作るには、学習者が知っているトピックを選ぶようにして、できるだけ日本語以外の知識がテストの回答に影響しないように注意しましょう。

次の例題2は、形容詞の過去形を学習した後で、それを使うことができるかを確認するために作られた問題です。この問題ができたなら、形容詞を使って、友人に感想や様子を言うことができる、と判断していいでしょうか。

例題 2 次は、旅行について話している会話文です。完成させてください。

A: 旅行はどうでしたか。

B: (いそがしい⇒ )ですが、(おもしろい⇒ )です。

A: そうでしたか。(よい⇒ )ですね。

この問題では、形容詞の形を<sup>へんかん</sup>変換できるかどうかはわかりますが、実際の場面で意味と形を考え、自分の状況に合った形容詞を<sup>てきせつ</sup>適切に使えるかどうかは確認できません。実際に友人に様子や感想を言えるかどうかを確認するには、旅行した場所などについて「どうでしたか」「どんなところでしたか」という質問に対して、話したり書いたりするようなテストや、例題3の

ように友人同士という役割でロールプレイをするのが、妥当性のあるテストです。

### 例題3 ロールカード(母語)

友達に、あなたの旅行について質問されます。

旅行した場所がどうだったか、感想を話してください。

妥当性の高いテスト問題を作るには、問題の形式や内容が、測りたい力を測っているかどうかをよく確認することが大切です。

## 2) 信頼性

テストの結果が安定していることを**信頼性**と言います。いつ、だれがテストをしても、同じ結果が得られることを言います。

例題4は正しい答えを記入する問題ですが、答えには、「ふり始めた」「ふりだした」「びっくりした」など、いろいろ

例題4 \_\_\_\_に入る答えを書きなさい。

急に雨がふり\_\_\_\_\_。

考えられます。学習した項目だけに点数を与えるのか、間違っていないから点数を与えるのか、教師によって採点基準が変わってしまつては信頼性がありません。この場合は、例題4' のような**選択問題**にすると、誰が採点しても同じ点数になり、信頼性

例題4' \_\_\_\_に入る答えを選びなさい。

急に雨がふり\_\_\_\_\_。

①つづけた ②だした ③でた

は高くなると言えます。ただし選択問題では、本当に理解できていなくても、偶然正解する可能性があります。例題4は3分の1の確率で正解します。信頼性を高めるために、選択肢を増やしたり、問題の数を増やしたりして、テストの結果が偶然良くなったり悪くなったりしないように、注意しましょう。

例題5は、自分で文章を書く課題です。こうした自由作文や会話テストは、自分で文を書く力や、会話の力を測るために、妥当性の高いテスト問題だと言えます。しかし正しい答えは一つでは



ありませんから、採点者によって点数が変化したり、同じ教師が採点する場合でも、採点するときの気分で採点結果が変わったりすると、信頼性が無いテストになってしまいます。同じ力を持った学習者が受験したら、二人のテストの結果が同じになるよう、信頼性を高めることが必要です。そのためには、採点基準を用意したり、複数の教師で採点して結果について話し合ったりすることが有効です。

また、テスト問題のテーマや形式も、内容の理解や産出に影響しますので注意が必要です。よく知らないテーマや、慣れていない出題形式では答えにくい、ということに注意しましょう。特定の知識や経験がある学習者に有利になったり、不利になったりするとテストの信頼性が低くなります。テスト問題のテーマや形式に偏りがないかを確認することも大切です。



【タスク 3】 テストで、学習者が自分で文章を書く課題(作文)を出した場合、採点の信頼性を高めるために、どんな工夫が必要ですか。

### 3) 真正性<sup>しんせい</sup>

真正性とは現実に経験する可能性のことです。例えば会話テストでは、場面や役割の設定が学習者にとって身近で、経験する可能性があるものなら、真正性が高まります。また読解や聴解のテストでは、テキストとして学習者が実生活で実際に読んだり聞いたりする可能性が高い内容や種類を選ぶと真正性が高まります。

具体例を考えましょう。「友人に簡単な感想や様子を言える」という学習目標が達成できたか、学習項目としての形容詞について知識を身につけたかどうかを測る真正性のあるテスト問題とは、どのようなものでしょうか。

例題6は、形容詞を使うように指定され、どの程度形容詞が定着しているのか確認することができます。しかし、現実には、このような課題に出会うことはありませんから、真正性が低いと言えます。

#### 例題6

形容詞を5つ使って、あなたが住んでいる町の説明を書いてください。

一方例題7は、友人にメールを書くという、学習者に身近で関係がある設定で、形容詞を使う必要がある

#### 例題7

あなたが住んでいる町に関心を持っている友人からメールが来ました。町の様子をくわしく伝える返信メール<sup>へんしん</sup>を書いてください。



課題です。人に説明するということは、実際に出会う可能性があり、真正性が高い課題だと言えます。

真正性はなぜ大切でしょうか。真正性のあるテストの結果によって、学習者が現実場面での程度日本語を運用できるのか、予測することができるからです。また、真正性のあるテスト問題は、学習に良い波及効果<sup>はきゅうこうか</sup>を与えます。

【タスク 4】 動詞の過去形についての知識が身につけたかどうかを確認する、真正性が高いテスト問題を考えてください。

#### 4) 波及効果<sup>はきゅうこうか</sup>

波及効果とはテストや評価が学習に与える影響<sup>えいきょう</sup>のことです。学習者はテストの結果に大きな関心を持ち、良い結果が出るように努力することが多いでしょう。テストが真正性の高いものであれば、テストの準備のための学習が、現実の運用<sup>うんよう</sup>に役に立つ学習になります。そして、実際にテストで出された問題は「大事なポイント」として学習者に理解されます。

例えば、テストが会話力を測るパフォーマンスタスク中心だったら、会話をもっとがんばろうと思って学習する人が増えるでしょう。反対に、目標が会話力の育成<sup>いくせい</sup>であっても、テストで文法の問題が中心だったり、記号を選ぶ問題が中心だったりしたら、学習者は、それに合わせた学習をするようになるでしょう

良いテストを作るには、テストにどのような波及効果があるか、問題の形式や内容が学習に良い影響を与えるかを考える必要があります。

## 5) 実用性<sup>じつようせい</sup>

良いテストが作れたとしても、テストを作ることや、テストの実施が大変では、実用性がありません。教育現場の実情<sup>じつじょう</sup>に合わせたテストを行うことを忘れないようにしましょう。例えば、テストを作る時間と人とお金があるか、テストをする時間と場所があるか、会話テストであれば、テスターや採点者がいるかなどは必ず確認する必要があります。テストを計画する時から、実用性と他の4つの条件をチェックしておきます。

ただし実際には、すべての条件を満たすテストを作ることはなかなか難しいものです。自分の教育現場に応じて、どの条件を優先<sup>ゆうせん</sup>するのかを考えて、テストを作ることが大切です。

【タスク 5】 良いテストを作るための観点を下の表にまとめてください。下の選択肢から適

当な言葉を選び、空欄<sup>くうらん</sup>を補ってください。(2回使う言葉もあります)

観点	説明
妥当性	テストが( )を測っているか、ということ。 テストの( )に注意し、( )以外の知識を測らないようにすることが大切。
信頼性	テストの( )が安定していること。 選択式の問題では( )や( )の数を多くし、記述式の問題では( )を用意するなど工夫が必要。

真正性	テストの課題を( )する可能性があるか、ということ。
波及効果	テスト問題の( )が( )のこと。
実用性	テストを( )しやすいか、ということ。

<選択肢> 経験、測りたい力、内容や形式、実施、採点基準、  
学習に与える影響、結果、問題、選択肢、日本語

【タスク 6】 次の2つの問題は、「てもいいです」を使うことができるかどうかを確認するためのテスト問題です。良いテストを作るための5つの観点から、2つの問題を比べてください。

#### 問題 A

「てもいいです」を使った文を2つ書いてください。

#### 問題 B

美術館で写真を撮りたいです。  
美術館の人に許可をとってください。

### まとめ

到達度テストを作るには、まず学習目標・学習内容を確認することが必要です。良いテストを作る観点には①妥当性、②信頼性、③真正性、④波及効果、⑤実用性があります。こうした観点と、テストの目的を考え合わせて、何を優先してテストを作るかを考えることが大切です。

**■ このパートの参考文献と参考サイト**

- 国際交流基金(2011)『学習を評価する』(国際交流基金日本語教授法シリーズ  
12)ひつじ書房
- 近藤ブラウン妃美(2012)『日本語教師のための評価入門』くろしお出版
- 関正昭・平高史成編(2013)『テストを作る』スリーエーネットワーク
- 根岸雅史(2012)「学校のテストはどうあるべきか:テスト作りの慣習を疑う」『英語教育』10月増刊号 pp.6-7 大修館書店
- 「JF 日本語教育スタンダード」<https://www.jfstandard.jp.go.jp>
- 「みんなの教材サイト」<https://www.kyozai.jp.go.jp> (ログインが必要)

**■ タスクの答え**

【タスク 1】【タスク 2】 省略

【タスク 3】 採点基準を用意したり、採点者同士で採点結果を話し合ったりする。作文のテーマや出題形式が学習者に分かりやすいものかどうか確認する。

【タスク 4】 省略

## 【タスク 5】

観点	説明
妥当性	テストが(測りたい力)を測っているか、ということ。  テストの(内容や形式)に注意し、(日本語)以外の知識を測らないようにすることが大切。
信頼性	テストの(結果)が安定していること。  選択式の問題では(問題)や(選択肢)の数を多くし、記述式の問題では(採点基準)を用意するなど工夫が必要。
真正性	テストの課題を(経験)する可能性があるか、ということ。
波及効果	テスト問題の(内容や形式)が(学習に与える影響)のこと。
実用性	テストを(実施)しやすいか、ということ。

## 【タスク 6】

- ・妥当性=「てもいいです」が必要な場面で使えるかどうかを確認できるのが、妥当性がある問題です。問題 B の方が妥当性が高い問題です。
- ・信頼性=誰が採点しても安定した結果が得られるようにするには、どちらの問題も、採点基準が必要です。「写真をとってもいいですか」のように、形に誤りがある解答をどう採点するかについての基準を作り、信頼性を高める必要があります。

- ・真正性=現実の生活では、ある表現を使うことを指定されて文を作るということはありませんので問題 A は真正性が低いです。
- ・波及効果=問題 B は現実場面で使う可能性があり、真正性が高い問題なので、良い波及効果が期待できます。
- ・実用性=問題 A は、簡単に準備できテストしやすいですが、問題 B は口頭テストのため、実施にも採点にも時間がかかり、学校や教育現場の条件によっては、実用性が高いとは言えません。