

# **CAN-DO: LINKING TESTS TO FRAMEWORKS OF LANGUAGE PROFICIENCY**

**Jakarta, 21 March 2019**

**Sheryl Cooke 柯雪柔**

***East Asia Assessment Solutions Team***

***British Council***

# AIMS AND OVERVIEW

## Overview

- ▶ Purpose and role of standards and frameworks
- ▶ Linking:
  - ▶ How linking is done: a case study
  - ▶ Practical application of linking procedures

# WHAT ARE STANDARDS AND FRAMEWORKS?

- **Not** tests
- Describe language proficiency on a common scale
- Often use a set of ‘can-do’ statements
- Reference document
- ‘Policy instrument’ – Council of Europe Companion Volume, 2018
- Standards vs frameworks

The Common European Framework provides a **common** basis for the **elaboration** of language **syllabuses**, **curriculum guidelines**, **examinations**, **textbooks**, etc. across Europe. – CEFR, 2001, p. 2

# THE PURPOSE OF STANDARDS

- A **common reference** point – transparency and coherence
- **Measurement across** schools, countries, languages
- **Benchmark** for situating teaching, learning and assessment
- **Drives** learning and teaching

Fundamentally, the CEFR is a tool to assist the planning of curricula, courses and examinations by working backwards from **what the users/learners need to be able to do in the language.**

CV, 2018, p. 26

The stated aims of the CEFR are to:

- ▶ promote and facilitate **co-operation** among educational institutions in different countries;
- ▶ provide a sound basis for the **mutual recognition** of language qualifications;
- ▶ assist learners, teachers, course designers, examining bodies and educational administrators to **situate and co-ordinate** their efforts. (CEFR: Section 1.4)

- CV, 2018. p. 26

The main functions of the CEFR are:

- (a) to provide a **metalanguage** for discussing the complexity of language proficiency and for reflecting on and communicating decisions on learning objectives and outcomes that are coherent and transparent, and
- (b) to provide inspiration for **curriculum** development and teacher education. -CV, 2018, p. 22

It aims to facilitate **transparency** and **coherence** between **curriculum, teaching and assessment** within an institution and transparency and coherence between institutions, educational sectors, regions and countries - CV, p. 25

The CEFR is being used not only to provide **transparency and clear reference points** for assessment purposes, but also increasingly to inform **curriculum reform and pedagogy**.  
- CV, p. 25

<https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>

# STANDARDS: GLOBAL VS LOCAL

## Council of Europe Framework of Reference (CEFR)

### Local vs Global

- Local use of language – context, variety
- Stronger link to the education system
- Different models of language and learning

CEFR-J

China Standards of English (CSE)

COMMON EUROPEAN FRAMEWORK  
OF REFERENCE FOR LANGUAGES:  
LEARNING, TEACHING, ASSESSMENT  
COMPANION VOLUME  
WITH NEW DESCRIPTORS



# THE NATURE OF FRAMEWORKS AND STANDARDS

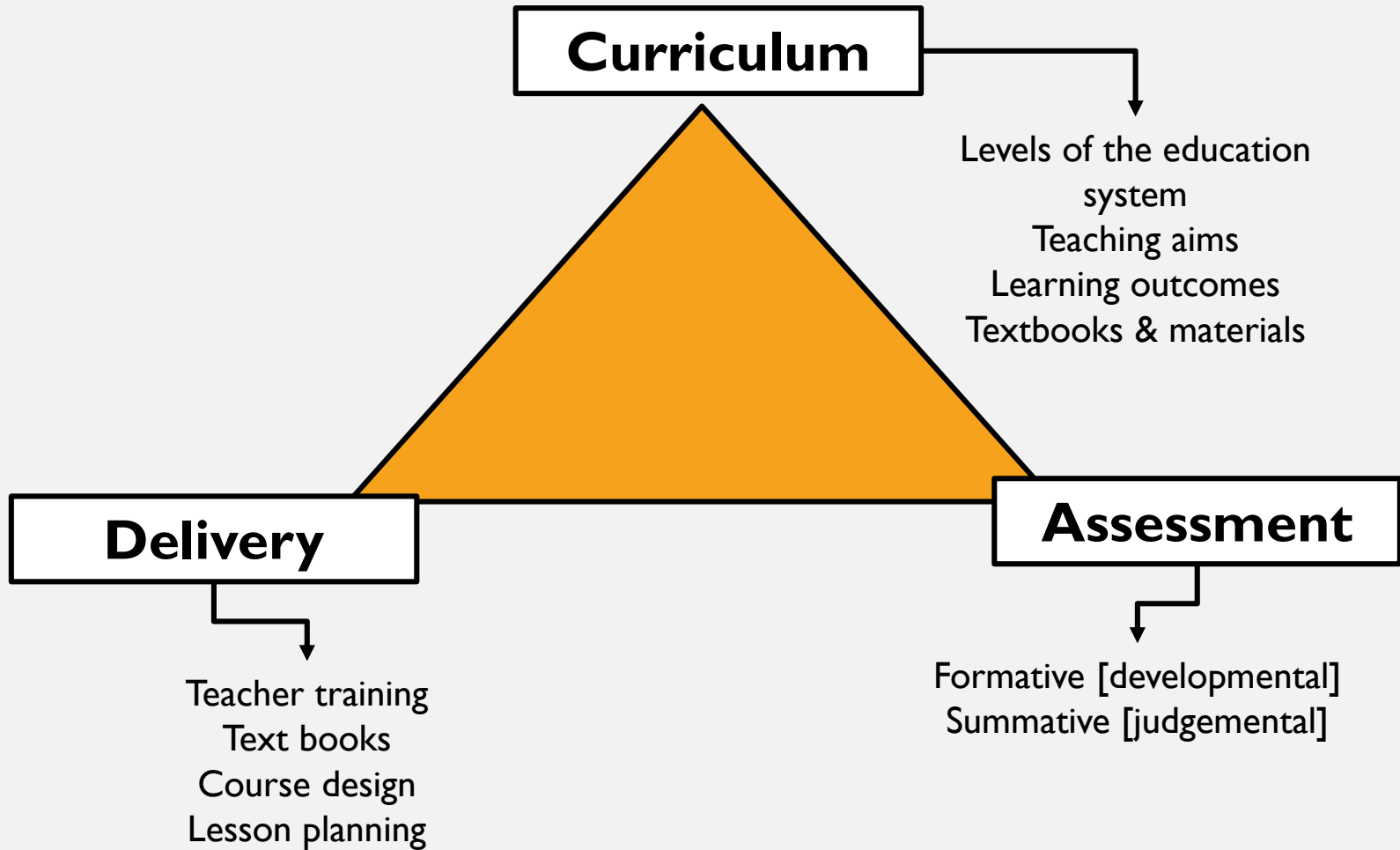
## What the CEFR does

- Proficiency perspective
- Real-world communicative needs
- Real-life tasks
- A tool to assist planning of
  - Curricula
  - Courses
  - Examinations

## What the CEFR does NOT do

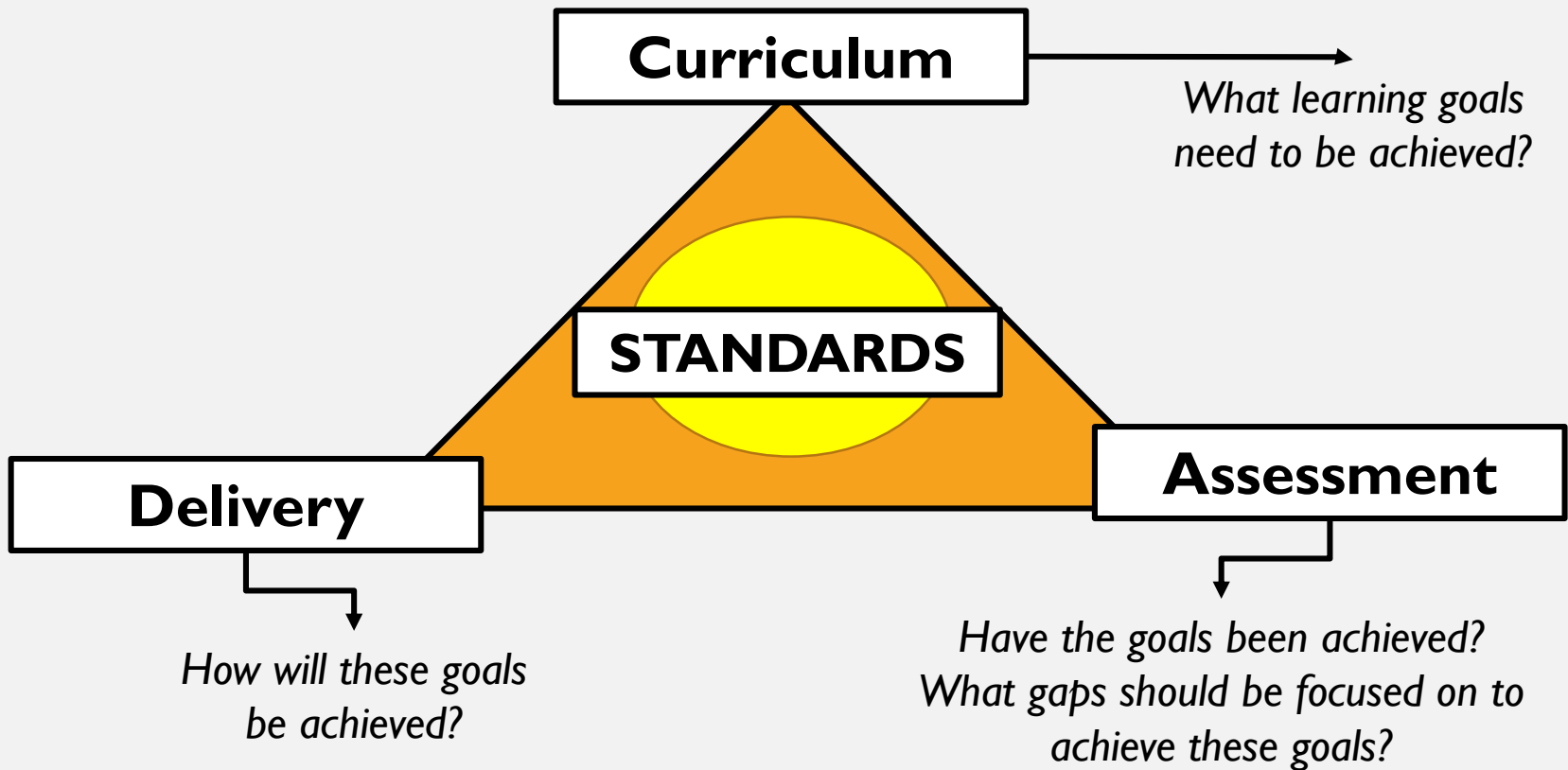
- Deficiency perspective
- Focus on linguistic knowledge rather than ability to use the language
- Prescribe to practitioners what to do or how to do it
- Set objectives or learning methods for learners

# THE LEARNING SYSTEM

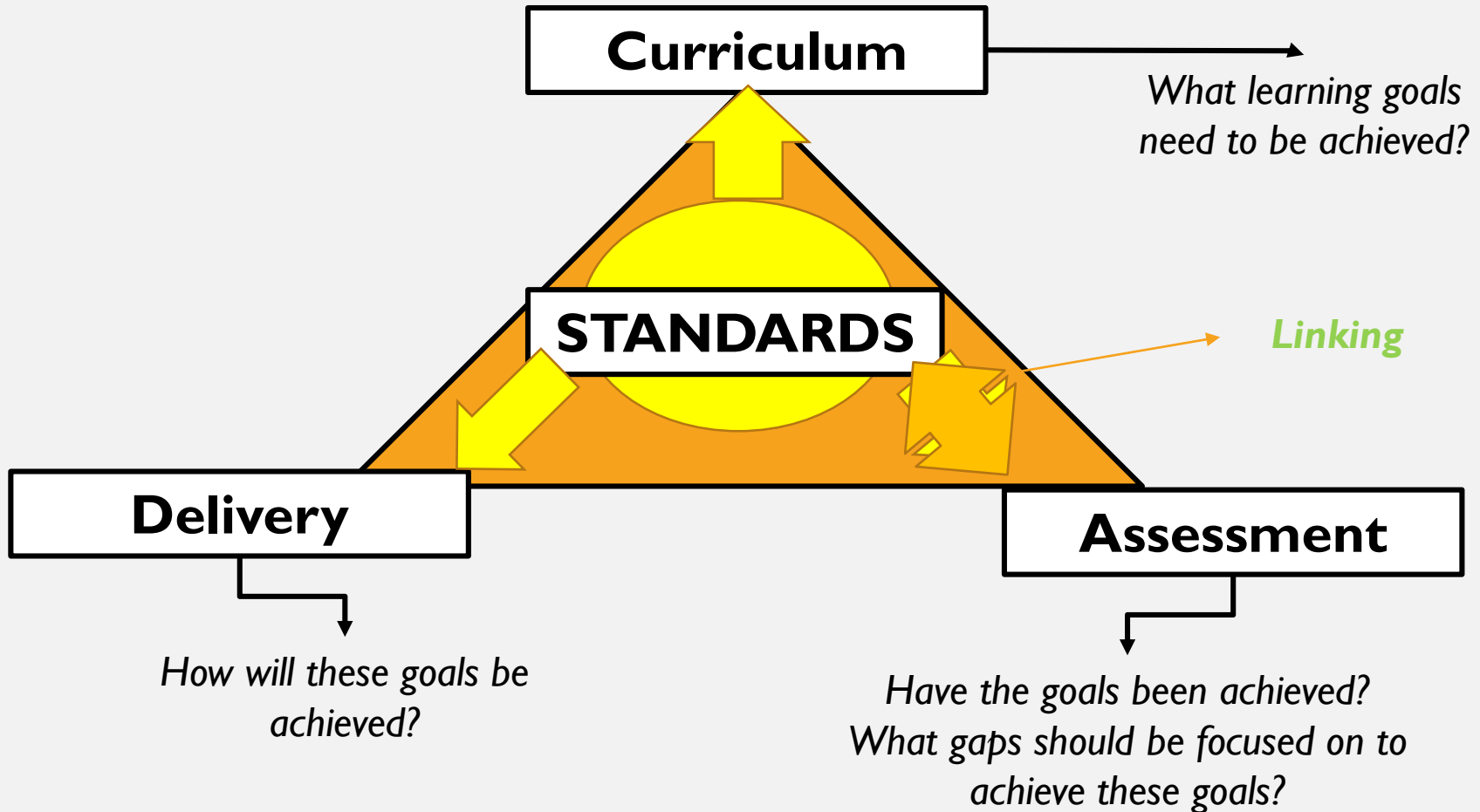




# THE LEARNING SYSTEM



# THE LEARNING SYSTEM



# WHAT IS STANDARD SETTING? (‘LINKING’)

- Allows tests to be **aligned to a framework** so that evidence-based judgements can be made about performance
- Allows organisations and institutions to **set cut scores** for the differentiation of pass/fail performances or for the selection of test-takers for a specific purpose

*The proper following of a prescribed, rational system of rules or procedures resulting in the assignment of a number to differentiate between two or more states or degrees of performance.*

- Cizek, 1993, p.10

*standard setting* refers to **the process of establishing one or more cut scores on examinations**. The cut scores divide the distribution of examinees' test performances into two or more categories.

- Cizek, 2007, p.5

The activity of **creating the boundaries of the performance categories** is one of the most important tasks in the test development, administration, and reporting process. That importance is due, in large measure, to the **consequences** that can be associated with the resulting classifications.

- Cizek, 2007, p. 6-7

# LINKING METHODOLOGIES

- Draw on expertise in the subject field
- Judgements about performance or difficulty
- Some Methodologies:
  - Analytical Judgement Method (AJM)
  - Basket Method
  - Angoff Methods
  - Bookmark Method
  - Body of Work Method
  - Constrasting Groups Method

# THE SCIENCE BEHIND IT

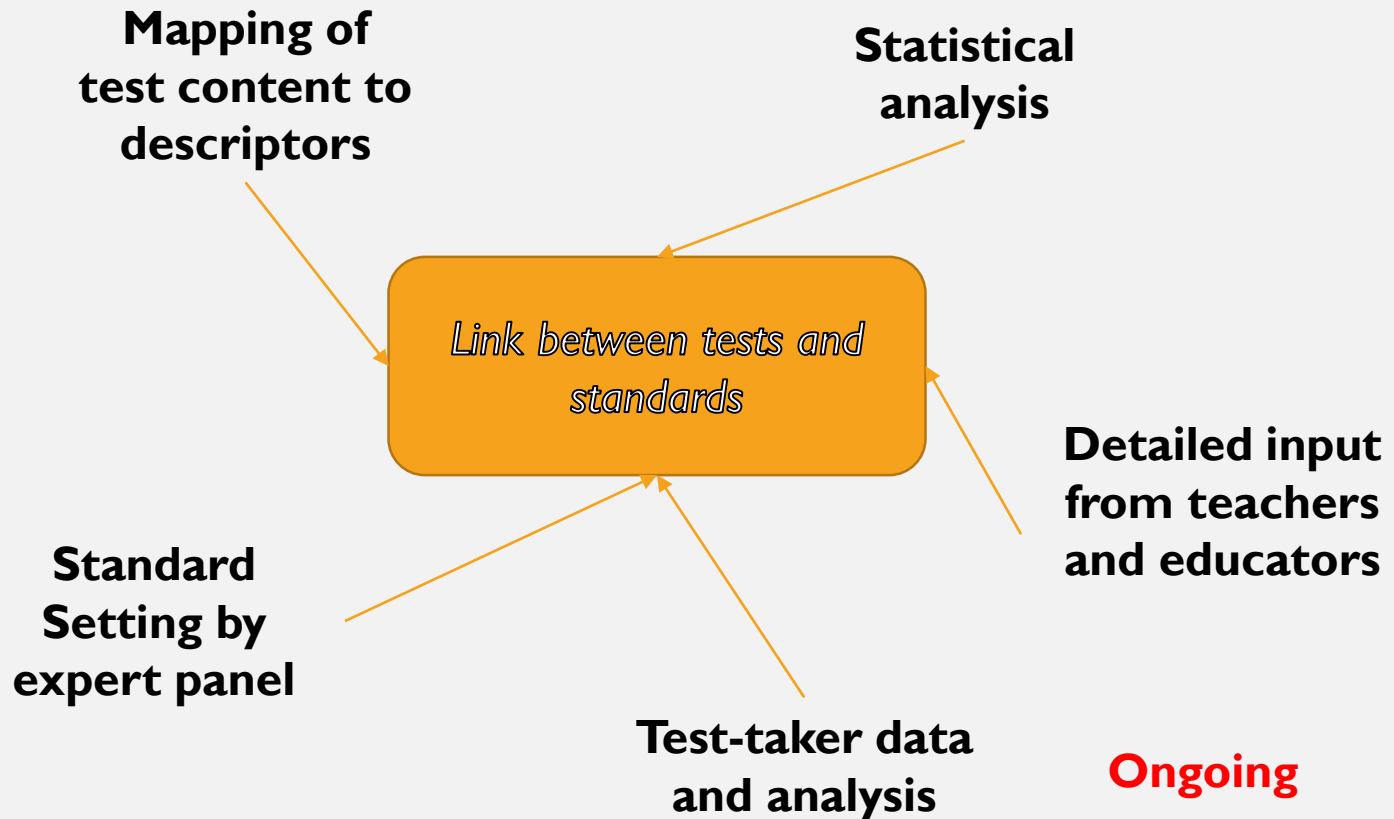
- Dunlea, J. (2014). *Investigating the relationship between empirical task difficulty, textual features and CEFR levels*. Paper presented at the 11<sup>th</sup> EALTA conference, University of Warwick, UK.
- Dunlea, J. (2016). *Validating a set of Japanese EFL proficiency tests; Demonstrating locally designed tests meet international standards*. Unpublished PhD thesis: University of Bedfordshire, Bedfordshire.
- Dunlea, J., & Figueras, N. (2012). Replicating results from a CEFR test comparison project across continents. In D. Tsagari and I. Csepes (eds.), *Collaboration in language testing and assessment* (pp. 31-45). New York: Peter Lang
- O'Sullivan, B. (2010). The City and Guilds Communicator examination linking project: a brief overview with reflections on the process. In W. Martyniuk (Ed.), *Aligning tests with the CEFR: Reflections on using the Council of Europe's draft manual*. Cambridge: Cambridge University Press.
- O'Sullivan, B. (2015b). *Linking the Aptis reporting scales to the CEFR*. Aptis Technical Report, TR/2015/003. London: British Council.
- Wu, R., Yeh, H., Dunlea, J. & Spiby, R. (2016). Aptis-GEPT comparison study: Looking at two tests from multiple perspectives using the socio-cognitive model. British Council Validations Series VS/2016/002. London: British Council.



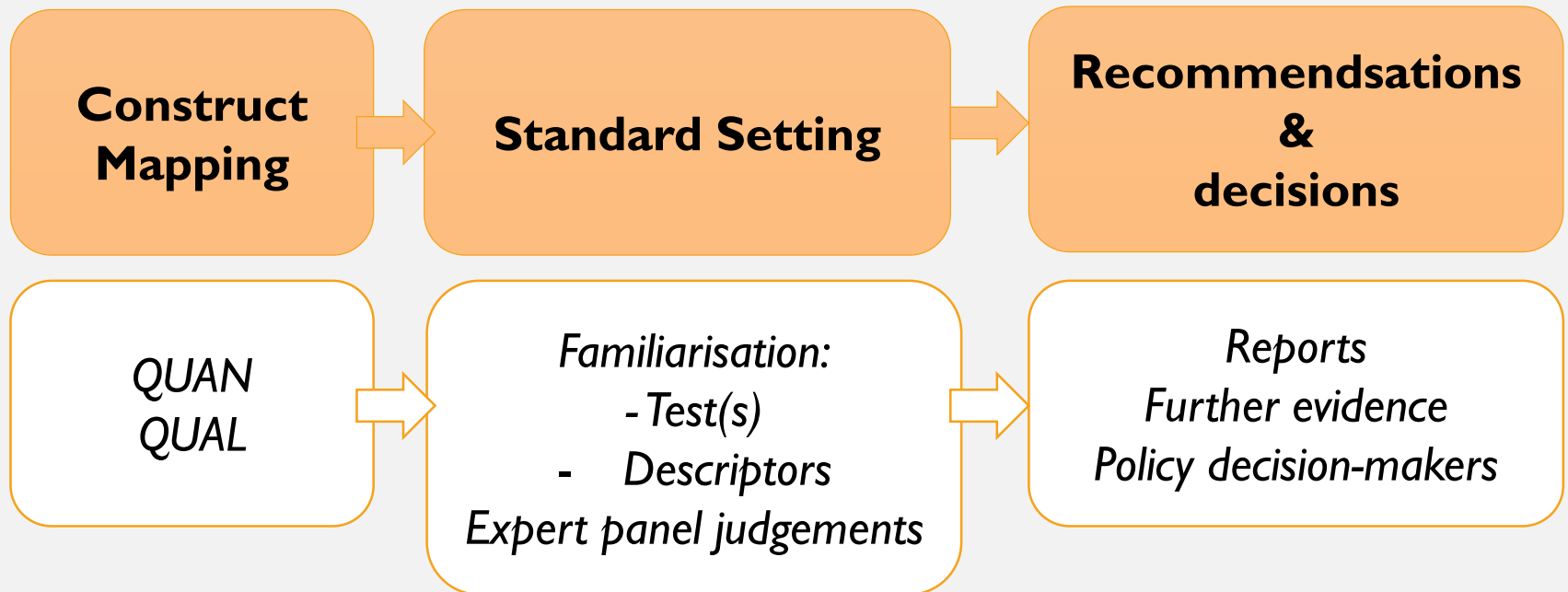
## IN PRACTICE

- Describes how some of the linking processes used in linking Aptis and IELTS to the CSE can be adapted for use in the classroom, e.g. to link formative classroom tests to the CSE
- Understanding and skills can be integrated into teaching practice

# LINKING AS AN ARGUMENT



# STEPS IN THE LINKING PROCESS





# CONSTRUCT MAPPING



*Test A*



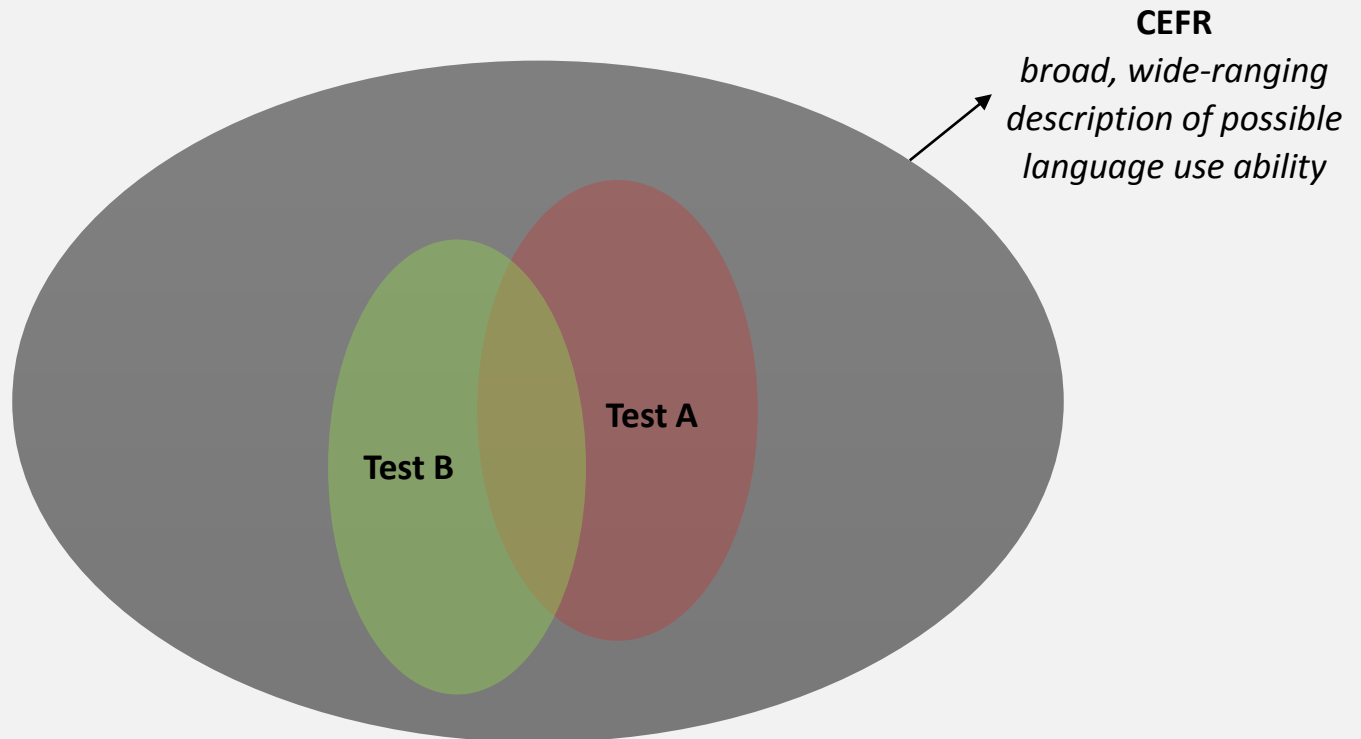
*Test B*



- *Fruit*
- *Shared characteristics*
  - *Not the same*

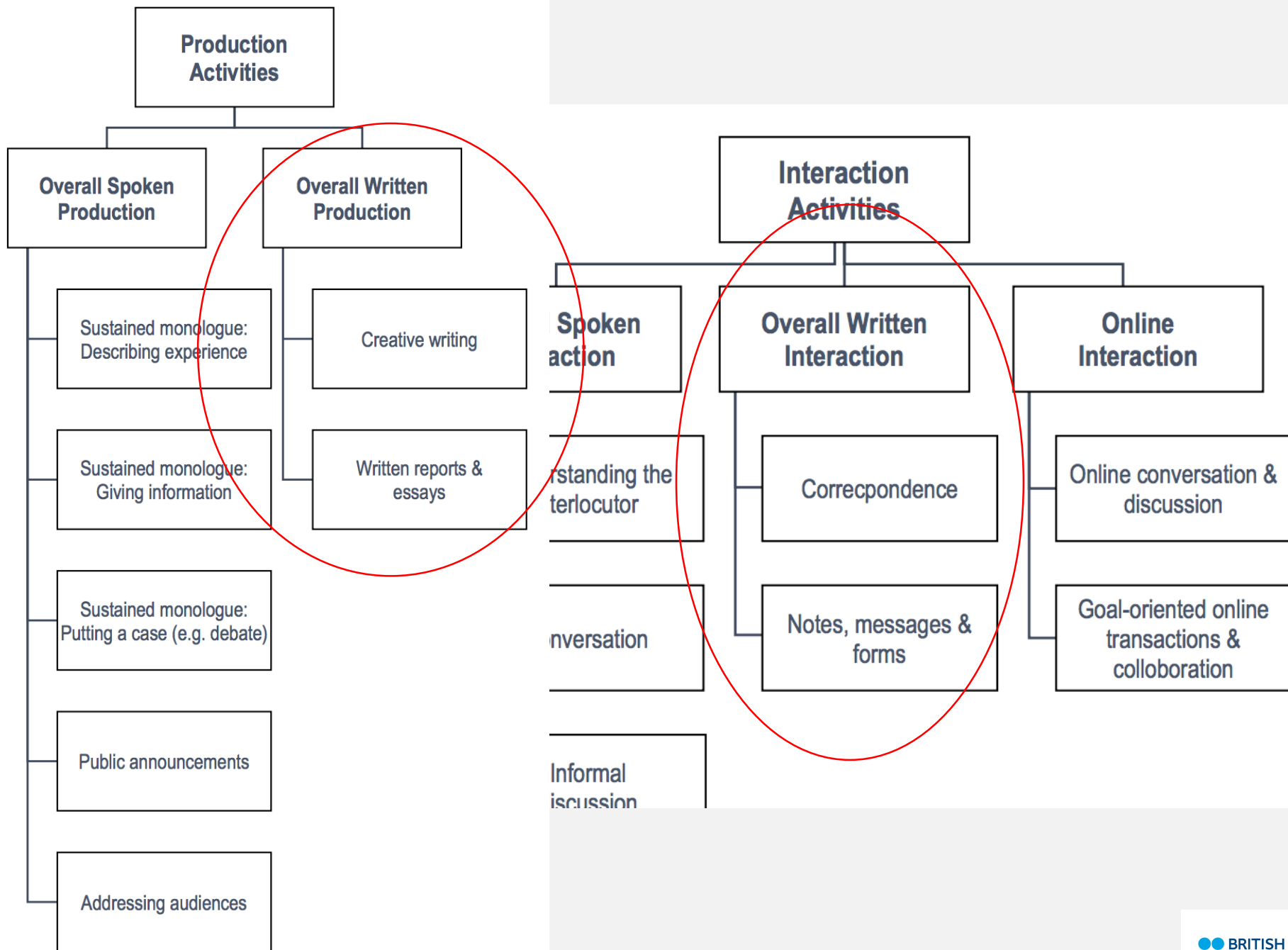
***Standards: descriptors  
of a wide range of  
language proficiency  
abilities***

# LINKING TESTS TO STANDARDS



# CONSTRUCT MAPPING

- Test / task / item
- What does it measure?
- How does it correspond to the (much bigger) construct described by the standards?





You are a member of a travel club. You received this e-mail from the club.

Dear Member,

We are writing to tell you that the famous travel writer Mr David Price will unfortunately not be able to attend our next club meeting. Although Mr Price will not be there to sign copies of his new book *Around The World In Eighty Ways*, members of the club will be able to buy a copy at the price of twenty five pounds. If you would like to reserve a copy of the book, please contact the club secretary.


Write an e-mail to your friend. Write about your feelings and what you think the club should do about the situation. Write about 50 words. You have 10 minutes.


Write an e-mail to the president of the club. Write about your feelings and what you think the club should do about the situation.


Write 120-150 words. You have 20 minutes.


***What aspects of the writing skill is the task focusing on?***


***How does the task relate to the framework?***


Can write straightforward, detailed descriptions on a range of familiar subjects within his/her field of interest. 


Can write accounts of experiences, describing feelings and reactions in simple connected text. 


Can write a description of an event, a recent trip – real or imagined. 


Can narrate a story. 


Can write short, simple essays on topics of interest. 


Can summarise, report and give his/her opinion about accumulated factual information on familiar routine and non-routine matters within his/her field with some confidence. 


Can express news and views effectively in writing, and relate to those of others. 


Can convey information and ideas on abstract as well as concrete topics, check information and ask about or explain problems with reasonable precision. 

Can write personal letters and notes asking for or conveying simple information of immediate relevance, getting across the point he/she feels to be important. 

Can write personal letters giving news and expressing thoughts about abstract or cultural topics such as music, films. 

Can write personal letters describing experiences, feelings and events in some detail. 

Can take messages communicating enquiries, explaining problems. 

Can write notes conveying simple information of immediate relevance to friends, service people, teachers and others who feature in his/her everyday life, getting across comprehensibly the points he/she feels are important. 

Language expression	Oral expression	Oral description
		Oral narration
		Oral exposition
		Oral argumentation
		Oral instruction
		Oral interaction
	Written expression	Written description
		Written narration
		Written exposition
		Written argumentation
		Written instruction
		Written interaction

- Can briefly state opinions on common topics. ✓
- Can briefly express how he/she feels after taking part in certain social activities. ✓
- Can briefly explain familiar but abstract concepts (e.g. friendship or happiness). ✗

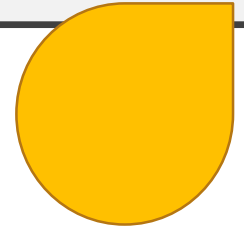
- Can briefly describe steps involved in carrying out routine activities. ✗
- Can write a clear travel itinerary for outdoor group activities. ✗

- Can make suggestions on how to solve problems in his/her life or study. ✓
- Can clearly explain the pros and cons of a particular action. ?
- Can persuade others to accept his/her points of view. ✓

- Can write letters or emails describing his/her personal information, interests, hobbies, and campus life. ✓
- Can write letters or emails to briefly describe familiar places (e.g. scenic spots). ✗
- Can briefly outline Chinese culture (e.g. traditional festivals and customs). ✗
- Can briefly describe his/her study plans, learning experiences, etc. in different contexts, including social media. ✗



# **IN PRACTICE: CONSTRUCT MAPPING**



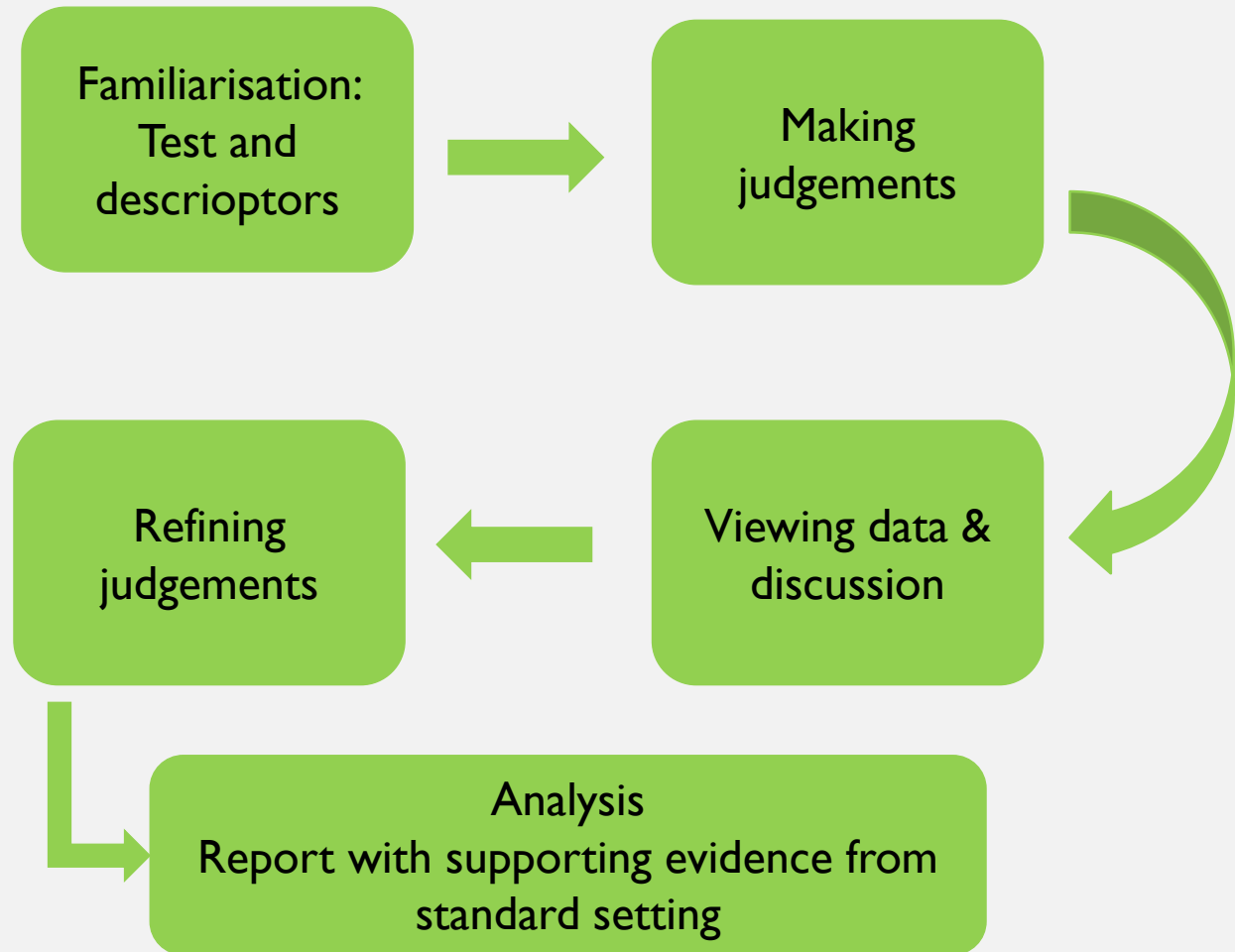
- How do the tasks and items on different exams relate to the framework? What ‘can-do’ statements do they link to?
- What materials can be included in curricula and what activities can be included by teachers to develop these language abilities?

# STANDARD SETTING

- A panel of experts
- Working to reach a consensus
- Statistical analysis of judgements
- Different methodologies used: Basket method, Modified Angoff, Analytical Judgement Method (AJM)



# STANDARD SETTING



# FAMILIARISATION

- In this step, an expert panel is asked to familiarise themselves with the tests that are going to be linked as well as the descriptors in a particular framework.
- This is usually done individually followed by group activities.
- This step is important because an in-depth understanding of the descriptors and the differences in levels is crucial for the following judgement steps in the process of linking.
- Clear understanding of the texts and which area of the descriptors they link to is also important.

# FAMILIARISATION: EXAMPLE ACTIVITIES

- Self-assessment
- Identifying higher vs lower descriptors
- Identifying which level descriptors relate to
- Ordering of descriptors from a set of scales
- Categorising descriptors into different scales (e.g. overall, correspondence, etc.)

# IN PRACTICE: FAMILIARISATION

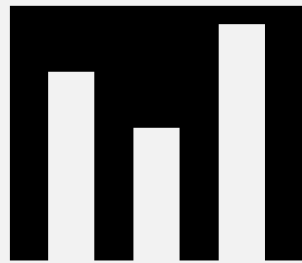


- Teacher training activities
- Using descriptors as diagnostic activities: self-diagnosis (learners) or teacher diagnosis of learners
- Relating test tasks and classroom activities to descriptors

# STANDARD SETTING



***Individual  
judgements***



***Sharing of group  
statistics***



***Group discussion: Why  
did you put it into that  
level?***



***Individual  
judgements***

# STANDARD SETTING METHODOLOGY

Receptive  
Skills

- Basket Method



- Angoff Method



Productive  
Skills


- Analytical Judgement Method




# THE BASKET METHOD


- A range-finding activity
- Conceptually more straightforward than the following method
- Provides correlating data


## Reading part 3

 00 : 25  
Hours Mins

Hide time

 **BRITISH COUNCIL**

 **Aptis**  
English for Aptis

Screen 3 of 4 

Read the text and complete each gap with a word from the list at the bottom of the page.


**Galileo Galilei**


Often **referred** to as 'the father of modern physics', Galileo Galilei was born in Pisa, Italy, in 1564, the son of a mathematician and musician. He attended university in Pisa but had to leave due to a **lack** of funds, and later taught sciences at the University of Padua. It was **during** his time there that Galileo did a large number of **experiments**, the most famous involving dropping balls of different sizes from **various** heights to determine the law of acceleration of falling bodies. Indeed, he is credited with several important scientific **discoveries**, and is still considered a great genius. Unfortunately, however, he died in prison in 1642, whilst **serving** a life sentence for publishing works suggesting that the earth moved **around** the sun – something that went against accepted thinking at the time.


taking

lot

at

 **Back**

**Flag** 

**Next** 

# THE BASKET METHOD

*At which CSE level can a candidate FIRST answer this item correctly?*



A0



A1



A2



B1



B2



C1



C2

*Decide which basket each item belongs in.*

# MODIFIED ANGOFF METHOD

## *The “Minimally Competent Candidate”*

A ‘minimally competent candidate’ is a test candidate that just falls inside a proficiency level. A ‘minimally competent candidate’ at B1 will have just enough English proficiency to receive a B1 certification.

- More refined approach
- Provides more granular statistics
- Conceptually more challenging

# MODIFIED ANGOFF METHOD



***Minimally Competent  
Candidate***



# MODIFIED ANGOFF METHOD


- Imagine a minimally competent examinee at B1 level.
- Imagine 100 candidates at the same level.



- Review each item.
- *How many (%) minimally competent examinees at the target level will answer the item correctly?*
- Only use 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 (%)

## Reading part 2

Screen **2** of **4** 

Order the sentences below to make a story. The first one is done for you.

### My First Day At Work

My first day at work didn't go according to plan.

I arrived early to make a good impression, but when I got there the lift was out of order.


The office is on the 22nd floor but I decided to use the stairs.


While walking up them I fell over and hurt my knee.


By the time I got to the office I couldn't walk anymore.

My new boss told me to go to the hospital to make sure the problem wasn't serious.

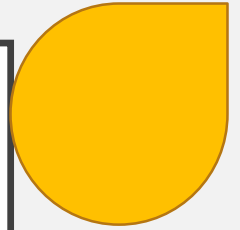
I didn't even get to see my new desk before I left.

 Back

Flag 

Next 

## IN PRACTICE: BASKET AND MAM



- ▶ Aligning institution or classroom tests to a framework
- ▶ Expert panel of teachers, for example
- ▶ Use the Basket Method to range-find or for low-stakes classroom tests
- ▶ If necessary, follow this up with a more refined approach such as the Modified Angoff Method



# LINKING THE PRODUCTIVE SKILLS

- Analytical Judgement Method
- Expert panel
- Examples of candidate responses
- Put the samples into levels – similar to the Basket Method

## IN PRACTICE: AJM



- Gather a range of student responses and share these between an expert panel (at least three teachers)
- Individually, rank the responses, giving them each a CSE level
- Come together to discuss; explain why you have ranked them in the way you have – always refer to the descriptors!
- Select responses that everyone agrees are characteristic of a particular CSE level
- In writing, describe the characteristics of the responses you have selected, being sure to refer back to key phrases in the CSE descriptors

# ASSESSMENT RESEARCH GROUP

**Technical Report**  
Aptis Test Development Approach  
TR/2015/001

## Aptis Test Development Approach

The Aptis test system is designed to offer users an alternative to currently available high-stakes certificated examinations.

**Technical Report**  
Linking the Aptis Reporting  
Scales to the CEFR  
TR/2015/003

## Linking the Aptis Reporting Scales to the CEFR

The Aptis development project marked a new era for the British Council, even though it had been involved in a number of test development projects in the past.

**Technical Report**  
Aptis General Technical Manual  
Version 1.0  
TR/2015/005

Barry O'Sullivan, British Council  
Jamie Dunlea, British Council

## Aptis General Technical Manual Version 1.0

TR/2015/005 This manual describes the content and technical properties of Aptis General, the standard English language assessment product offered within the Aptis test system.

**THANK YOU!**

**East Asia Assessment Solutions Team**

[eaastenquiries@britishcouncil.org](mailto:eaastenquiries@britishcouncil.org)

[www.britishcouncil.cn/en/east-people](http://www.britishcouncil.cn/en/east-people)

[sheryl.cooke@britishcouncil.org.cn](mailto:sheryl.cooke@britishcouncil.org.cn)

# USING THE CEFR ILLUSTRATIVE DESCRIPTORS

The main function of descriptors is to help

- **align curriculum, teaching and assessment.** Educators can select CEFR descriptors according to their relevance to the particular context, adapting them in the process if necessary. In this way descriptors can provide a detailed, flexible resource for:
  - ▶ relating learning aims to real world language use, thus giving a framework to action-oriented learning;
  - ▶ providing transparent ‘signposting’ to learners, parents, sponsors;
  - ▶ offering a ‘menu’ to negotiate priorities with adult learners in a process of ongoing needs analysis;
  - ▶ suggesting classroom tasks to teachers, usually tasks that will involve activities described in several descriptors;
  - ▶ introducing criterion-referenced assessment with the criteria relating to an external framework (here the CEFR).

# CHINA STANDARDS OF ENGLISH

- In September 2014, the State Council of China declared the “construction of the **assessment system** of foreign language proficiency” an imperative task. The Ministry of Education appointed the National Education Examinations Authority (NEEA) responsible for the assignment.
- Developing China’s Standards of English (CSE), is a fundamental task in constructing a complete nationwide assessment system.
- CSE **describes** what Chinese learners and users of English **can do with the target language**. It defines language proficiency levels comprehensively by specific, accurate and easy-to-understand descriptors in terms of language communicative functions and language skills.
- CSE serves as one of the fundamental strategies to promote English learning, teaching and assessment in China.

## CSE overall listening comprehension

---

- Can understand argumentation on abstract topics (e.g. politics, economy, history, culture); and evaluate the speakers' opinions and stance.

- Can understand academic discourse (e.g. lectures, operation instructions) related to one's own field; and comprehend main ideas and supporting details.

## CSE overall listening comprehension

---

### CSE 7

- Can understand argumentation on abstract topics (e.g. politics, economy, history, culture); and evaluate the speakers' opinions and stance.

### CSE 8

- Can understand academic discourse (e.g. lectures, operation instructions) related to one's own field; and comprehend main ideas and supporting details.



# Oral Description

---

- Can understand descriptions of certain places when delivered at a normal speed and grasp geographical features.

- Can obtain key information from descriptions of large-scale activities (e.g. festival celebrations, sports events), when delivered at a normal speed.

- Can understand descriptions of natural landscapes (e.g. mountains, rivers) regardless of speech rate; and summarise their main features.

# Oral Description

---

## CSE 6

- Can understand descriptions of certain places when delivered at a normal speed and grasp geographical features.

## CSE 5

- Can obtain key information from descriptions of large-scale activities (e.g. festival celebrations, sports events), when delivered at a normal speed.

## CSE 7

- Can understand descriptions of natural landscapes (e.g. mountains, rivers) regardless of speech rate; and summarise their main features.

# Comprehending Oral Argumentation

---

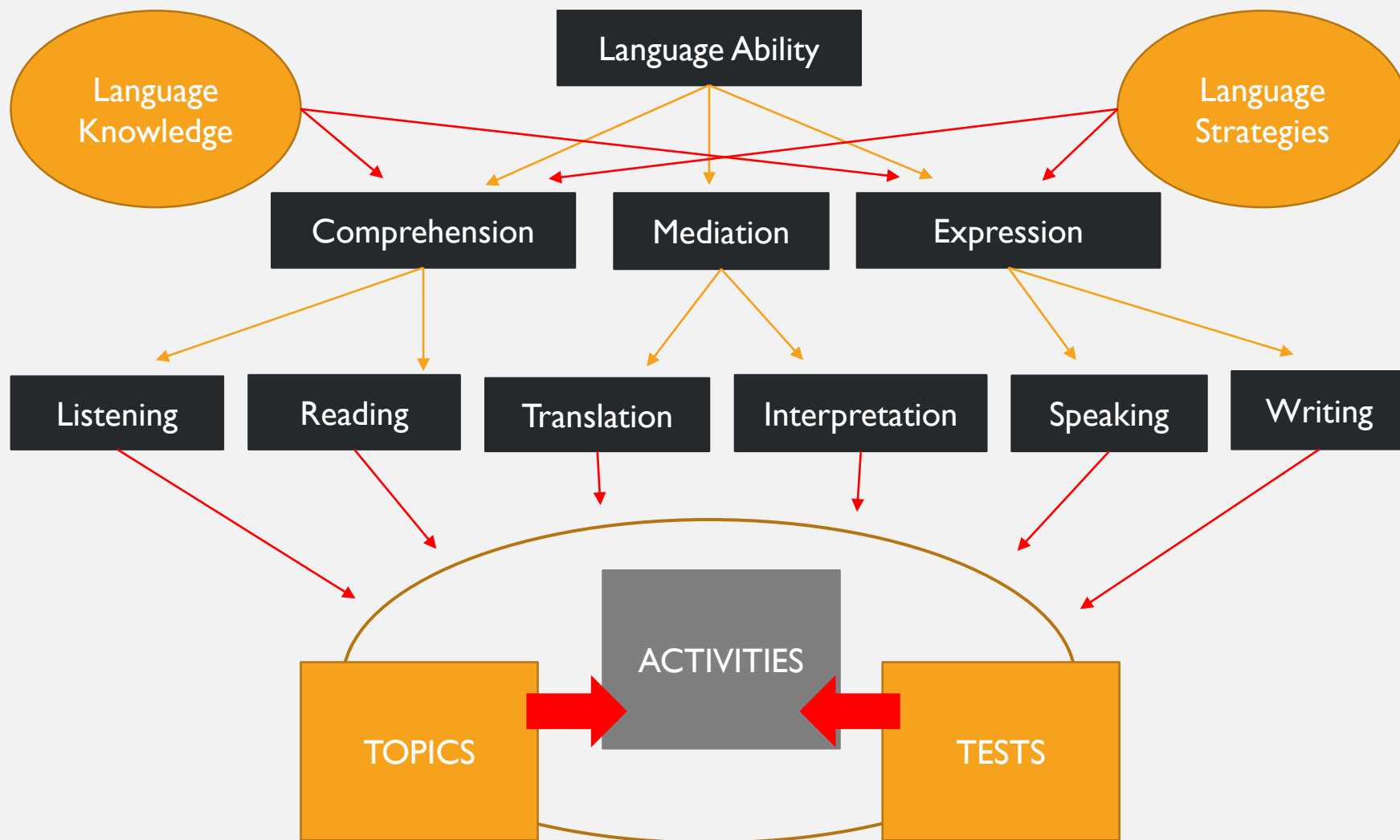
	A	Can understand discussions on social issues when delivered at a normal speed and evaluate the logic of argumentation.
	B	Can understand highly-informative lectures or audio-taped/ video-taped talks and summarise key points and opinions.
	C	Can understand academic conference presentations or debates in one's own field and evaluate speakers' main points.
	D	Can understand debates on familiar topics articulated clearly and grasp the main arguments and supporting evidence on both sides.
	E	Can understand academic lectures containing technical terms related to one's own field and comprehend the main content.
	F	Can understand short argumentation on familiar topics that is delivered in simple language at a slow but natural speed; and grasp the main idea.

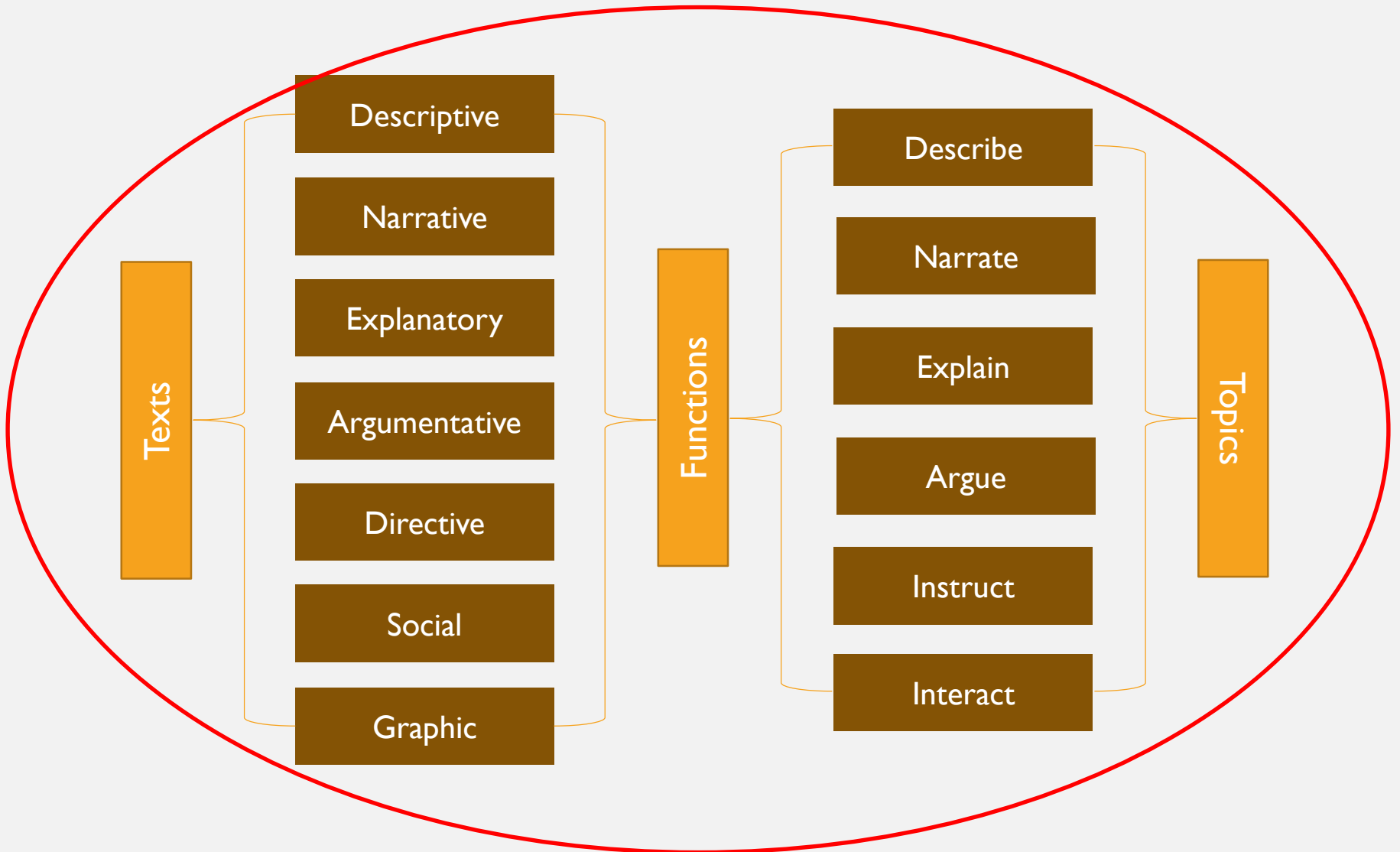
## Comprehending Oral Argumentation

---

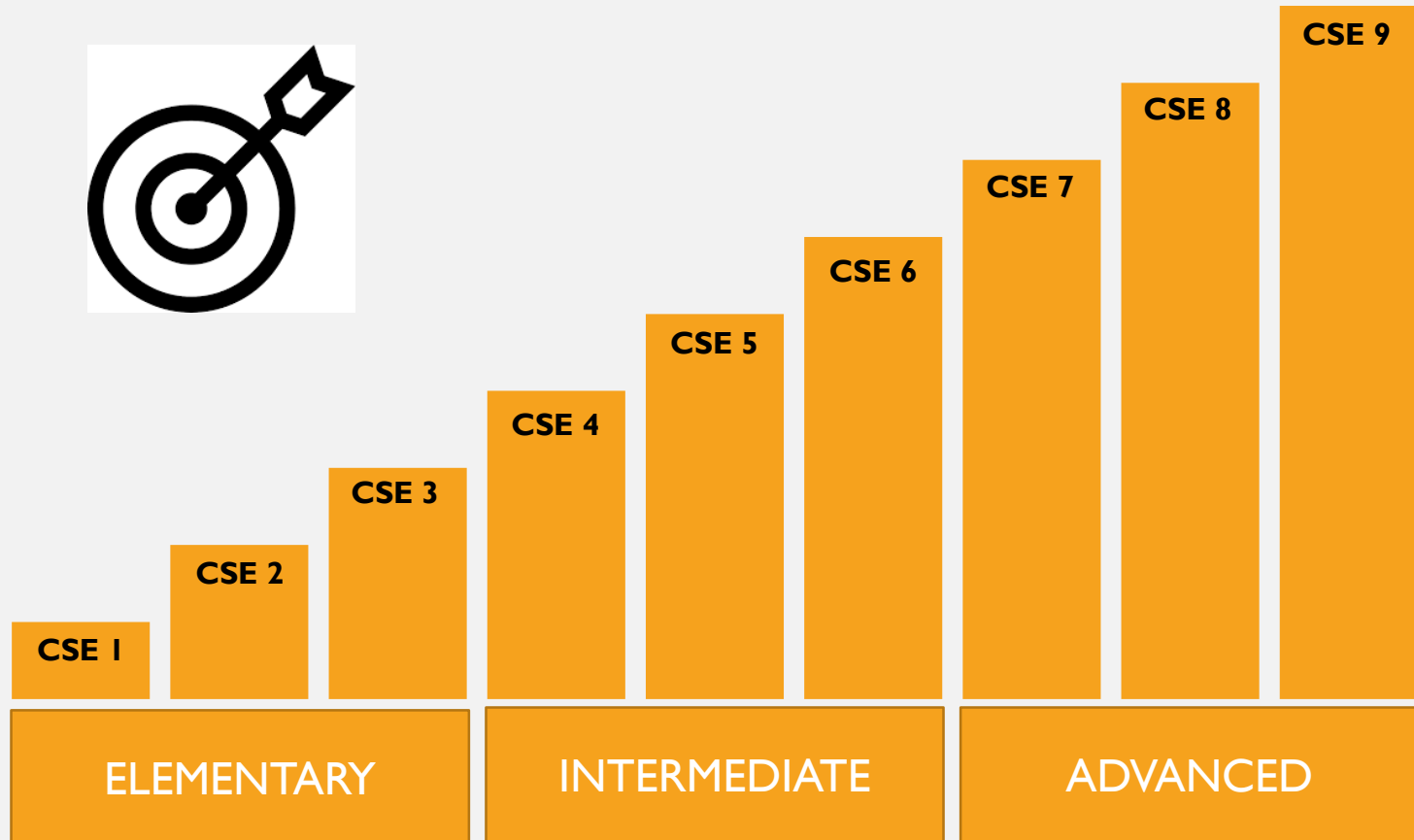
8	E	Can understand academic lectures containing technical terms related to one's own field and comprehend the main content.
7	C	Can understand academic conference presentations or debates in one's own field and evaluate speakers' main points.
6	B	Can understand highly-informative lectures or audio-taped/ video-taped talks and summarise key points and opinions.
5	A	Can understand discussions on social issues when delivered at a normal speed and evaluate the logic of argumentation.
4	D	Can understand debates on familiar topics articulated clearly and grasp the main arguments and supporting evidence on both sides.
3	F	Can understand short argumentation on familiar topics that is delivered in simple language at a slow but natural speed; and grasp the main idea.

# HOW CSE EVALUATES LANGUAGE





# CSE PROFICIENCY LEVELS



## NOTES ON CSE

- CSE is not intended to be enforced on any particular educational stage.
- CSE is not a teaching curriculum, no exam syllabus.
- CSE attempts to display the key features of English proficiency at each level rather than list all the language activities.
- CSE's revision and validation is an ongoing process, similar to CEFR as language and teaching is constantly evolving.
- Every user of CSE will contribute to its development.