



# 素养教育背景下的语言测评：变与不变

罗少茜  
北京师范大学



# Outline

- Education for a globalized world
- Framework for 21st Century Learning
  - English subject core competencies
    - China Standards of English
- Educational Assessment in a Changing World
  - A Chinese sample
    - **RCA + SBA**



# 1. Education for a globalized world

People need the ability to make sense of information, to tell the difference between what is important and what is unimportant, and above all to combine many bits of information into a broad picture of the world (Harari, 2018:261).

NEW YORK TIMES BESTSELLING AUTHOR  
OF SAPIENS AND HOMO DEUS

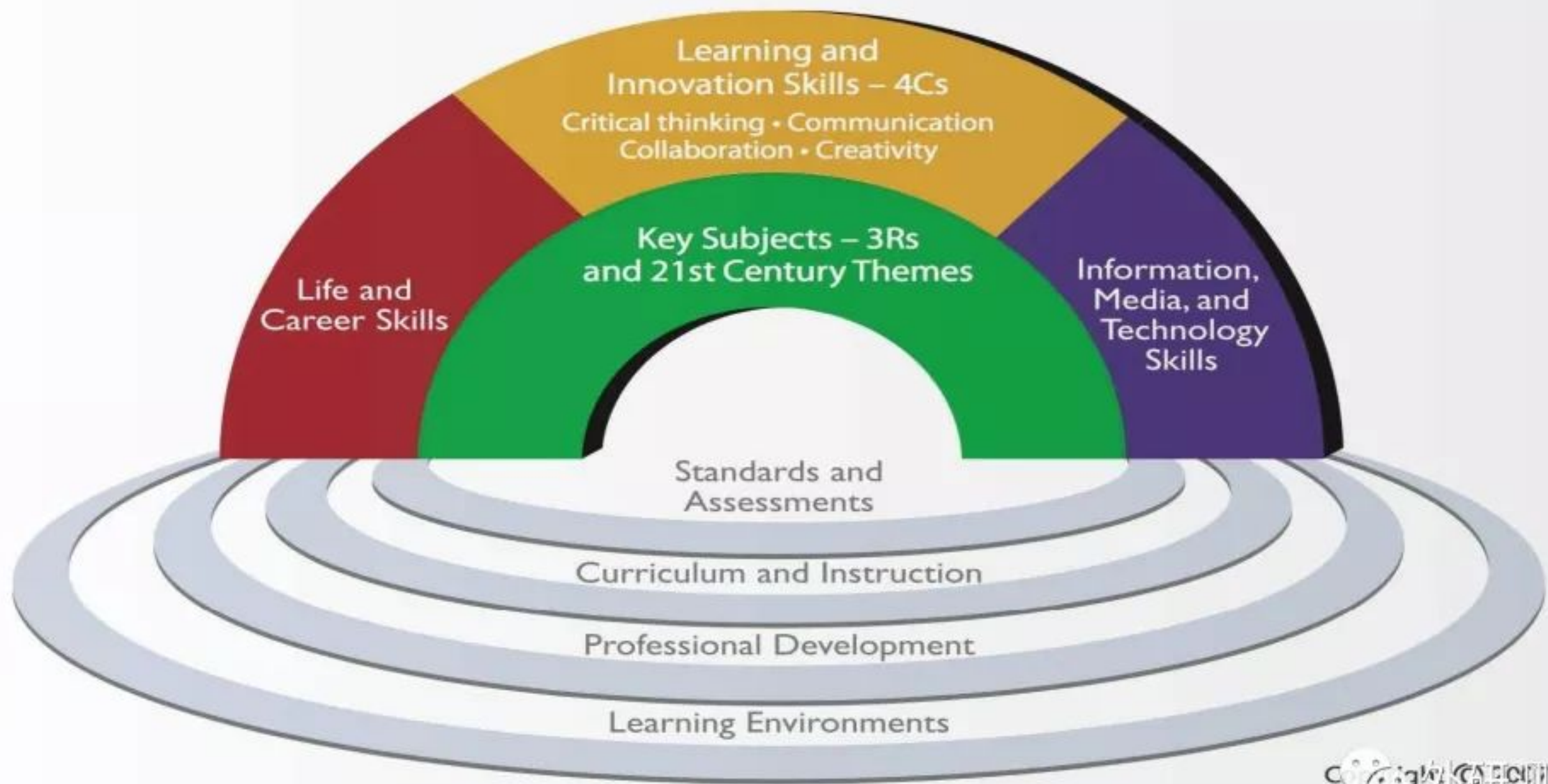
Yuval Noah  
Harari



21 Lessons  
for the  
21<sup>st</sup> Century



## 2. Framework for 21st Century Learning



### 3. 英语学科核心素养



语言能力



文化意识



思维品质



学习能力

# 英语学科核心素养定义及目标(1)

## • 语言能力

• 指在社会情境中，以听、说、读、看、写等方式理解和表达意义的能力，以及在学习和使用语言的过程中形成的语言意识和语感。英语语言能力构成英语学科核心素养的基础要素。英语语言能力的提高蕴含文化意识、思维品质和学习能力的提升，有助于拓展国际视野和思维方式，开展跨文化交流。

## • 高中阶段目标

• 具有一定的语言**意识**和英语**语感**，在常见的具体语境中整合性地运用已有语言**知识**，理解口头和书面**语篇**所表达的**意义**。识别其恰当表意所采用的手段，有效地使用口语和书面表达意义和进行**人际交往**。

# 英语学科核心素养的构成及定义(2)

## • 文化意识

• 文化意识指对中外文化的理解和优秀文化的认同，是学生在全球化背景下表现出的文化认知、态度和行为取向。文化意识体现英语学科核心素养的价值取向。文化意识的培育有助于学生增强国家认同和家国情怀，坚定文化自信，树立人类命运共同体意识，学会做人做事，成长为有文明素养和社会责任感的人。

## • 目标

• 获得文化知识，理解文化内涵，比较文化异同，汲取文化精华，形成正确的价值观，坚定文化自信，形成自尊、自信、自强的良好品格，具有一定的跨文化沟通和传播中华文化的能力。

# 英语学科核心素养定义及目标(3)

## • 思维品质

- 思维品质指思维在逻辑性、批判性、创新性等方面所表现的能力和水平。思维品质体现英语学科核心素养的心智特征。思维品质的发展有助于提升学生分析和解决问题的能力，使他们从跨文化视角观察和认识世界，对事物做出正确的价值判断。

## • 目标

- 能**辨析**语言和文化中的具体现象，梳理、概括信息，**建构**新概念，分析、推断信息的逻辑关系，正确**评判**各种思想观点，创造性地表达自己的观点，具备初步运用英语进行**独立思考、创新思维**的能力。



# 英语学科核心素养定义及目标(4)

## • 学习能力

- 学习能力指学生积极运用和主动调适英语学习策略、拓宽英语学习渠道、努力提升英语学习效率的意识和能力。学习能力构成英语学科核心素养的发展条件。学习能力的培养有助于学生做好英语学习的自我管理，养成良好的学习习惯，拓宽学习渠道，提高学习效率。

## • 目标

- 进一步梳理正确的英语学习观，保持对英语学习的**兴趣**，具有明确的学习**目标**，能够多渠道**获取**英语学习资源，**有效规划**学习时间和学习任务，选择恰当的**策略和方法**，监控、评价、反思和调整自己的学习内容和进程，逐步提高使用英语学习其他学科知识的意识和能力。

## 4. 中国英语能力等级量表(CSE)

### • 语言能力 language ability

语言学习者和使用者**运用**自己的语言知识、非语言知识以及各种策略，参与特定**情境**下某一**话题**的语言活动时表现出来的**语言理解能力**和**语言表达能力**。



## 5. Educational Assessment: What to Watch in a Rapidly Changing World

- Be technology based
- Measure “new” constructs
- Be built from richer underlying models of cognition and learning
- Make greater use of more complex tasks
- Be “personalized”
- Attempt to improve learning
- Be better at accounting for context
- Be “embedded” and distributed across time
- Use automated scoring
- Incorporate new approaches to modeling and analysis
- Provide more effective reporting



## 5. 未来教育测量的特征

- 以技术为依托
- 测量“新”的构念
- 建立在更深层次的认知和学习模型的基础上
- 更充分利用复杂任务
- 更“个性化”
- 试图改善学习
- 更好地考虑学生的背景
- “嵌入”教学活动并分布在不同时间
- 采用自动计分
- 把新的探索方法整合到建模和分析中
- 提供更有效的测量报告

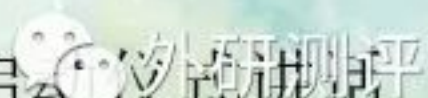
## 6. 课程标准与核心素养的关系

- Things to take into consideration:
  - 学科的思维方式
  - 核心素养的教与学
- 从语文学科的问题思考英语学科问题

# 语文教学的现实问题



郑国明，2018 北京十二中学科能力改进启动仪式讲话



基于语文传统单篇教学的反思

教学价值  
创作背景  
作者生平  
文体知识  
中心思想  
段落大意  
层次结构  
表现手法  
名言警句  
.....

过程：重教轻学

教学设计  
教学策略

学生

思维：程式套板

内容：肢解割裂

教师

单篇文本  
集体备课

能力：原地踏步

## 核心素养的评价

- 学生个体在解决真实的、复杂的、不确定性的现实问题过程中表现出来的综合性品质。
  - (1) 整合性：跨学科的过程与方法
  - (2) 潜在性：内在品质和特征



# 核心素养的评价

- 是学生个体在面对复杂的、不确定的**现实生活情境**
- 能够**综合运用**特定学习方式下所孕育出来的
  - （跨）学科**观念、思维模式和探究技能**
  - 结构化的（跨）**学科知识和技能**
  - 世界观、人生观和价值观在内的**动力系统**
- 分析情境、提出问题、解决问题、交流结果
- 过程中表现出来的**综合性品质**

## 核心素养的评价

- 40分钟
- Selecting a poem.
- Choosing a perspective to analyse it.
- Evaluate two comments on the poem.
- Recommend a poem to a senior high student.

## 7. 英语学科核心素养的测评

《课程标准》针对高中学业水平考试和高考提出了命题指导建议，包括命题原则、考查内容和形式。课程标准提出的对高中学业水平考试和高考的命题原则主要包括六个：

- (1) 根据高中英语课程的目标和理念确定命题导向和原则；
- (2) 全面考查英语学科核心素养；
- (3) 根据英语语言的实际使用情况命题；
- (4) 充分考虑学生生活经验和认知发展水平；
- (5) 确保试题的信度和效度；
- (6) 合理制定评分标准。

# 英语学科核心素养的测评



## 重点一

四者是相互关联的，因此评价也很难单独针对某一个方面进行评价，但可以通过分析学生的任务表现探索任务所能够探测到的学生素养。

# (1) RC: 阅读圈任务

- 也称文学圈任务 (literature circles) 是用于阅读教学的一项任务。
- 学生组成小组，分角色深入讨论学习所读的故事。
- 所分的角色可根据实际教学目标需求设置，也可按照Bookworms Club Bronze提供的六个阅读圈角色分工进行：
  - 讨论组长 (Discussion Leader)
  - 总结者 (Summarizer)
  - 文化收集者 (Culture collector)
  - 生活联系者 (Connector)
  - 单词大师 (Word master)
  - 文段解读者 (Passage Person)

- 阅读篇章：“Is your jar full?”
  - 一篇讲述人生哲理的叙事故事
  - 长度为420词。
- 任务要求：学生在规定时间40分钟内独立完成任务。
- 由于测试时间限制暂舍去“生活联系者”“篇章解读者”部分，即考生需要完成四个角色的任务：
  - ① 总结报告
  - ② 提问环节
  - ③ 词汇卡片
  - ④ 文化一角

## **Is your jar full?**

A Professor stood before his philosophy class and had some items in front of him. When the class began, wordlessly, he picked up a very large and empty glass jar and proceeded to fill it with golf balls. He then asked the students if the jar was full. They agreed that it was.

So the Professor then picked up a box of pebbles and poured them into the jar. He shook the jar lightly. The pebbles rolled into the open areas between the golf balls. He then asked the students again if the jar was full. They agreed it was. The Professor next picked up a box of sand and poured it into the jar. Of course, the sand filled up everything else. He asked once more if the jar was full. The students responded with a unanimous "Yes."

The Professor then produced two cans of beer from under the table and poured the entire contents into the jar, effectively filling the empty space between the sand. The students laughed.

"Now," said the Professor, as the laughter subsided, "I want you to recognize that this jar represents your life.

The golf balls are the important things - your family, your children, your health, your friends, your favorite passions - things that if everything else was lost and only they remained, your life would still be full.

The pebbles are the other things that matter like your job, your house, your car. The sand is everything else - the small stuff."

"If you put the sand into the jar first", he continued, "there is no room for the pebbles or the golf balls. The same goes for life. If you spend all your time and energy on the small stuff, you will never have room for the things that are important to you. Pay attention to the things that are critical to your happiness. Play with your children. Take time to get medical checkups. Take your partner out to dinner. Play another 18. There will always be time to clean the house, and fix the disposal. Take care of the golf balls first, the things that really matter. Set your priorities. The rest is just sand."

When he had finished, there was a profound silence. Then one of the students raised her hand and with a puzzled expression, inquired what the beer represented. The Professor smiled. "I'm glad you asked. It just goes to show you that no matter how full your life may seem, there's always room for a couple of beers."

1. Summarize the story in your own words.

My summary: A professor made a ~~ex~~ show about filled <sup>with</sup> the balls, and <sup>(a jar)</sup> he putted two cup of beers into the jar. Because He want to ~~ex~~ talk with student that we should find <sup>out</sup> the things is first or second in our life and we also have relaxing.

1. Summarize the story in your own words.

My summary: A professor stood before his philosophy class and had some items in front of him. So the professor then picked up. Then one of the students raised her hand and with a puzzled expression,

1. Summarize the story in your own words.

My summary: The ~~pa~~ <sup>professor</sup> told the students a significant <sup>theory</sup> ~~ther~~ <sup>represent</sup> the jar represents your life by doing a <sup>on their future life</sup> experiment, and <sup>may</sup> gives the student a profound influence.



3. Find at least FOUR words that are either

(1) new or difficult for you (if you can guess the meaning, explain your understanding)

Or (2) important in understanding the story and your reasons for choosing them.

My words: the meaning / the reasons

~~pet~~ pebble : small rock  
subside : sound  
items : things  
jaw : cup

My words: the meaning / the reasons

pebbles : a small solid items, that are smaller than golf balls and much <sup>bigger than sand</sup>  
subsided : stop, because professor cannot talk while they are laughing <sup>because it ~~was~~ <sup>should be</sup> laugh</sup>  
profound : constant, because students all thinking <sup>of</sup> the words <sup>of</sup> profes  
~~best~~ priorities : timetable, to ~~me~~ control the time wisely.

My words: the meaning / the reasons

unanimous : I guess the meaning is: without any doubts.  
subsided : I guess the meaning is: faded, stop.  
critical : it's important, cause. critical means: blame, comments or important.  
disposal : I guess the meaning is: the thing need to be fix.

5. Find at least ONE similarity and ONE difference between your culture and the culture found in this story.

My cultural collection:

Similarity: My family and children is the one of the important thing for the life. Many people just care of their child is take up all the time of their life.

Different: In China, the car isn't small things, you don't have car, then you don't have wife.

5. Find at least ONE similarity in this story.

My cultural collection:

Students should raise their hand before ask the question in order to show the respect. The difference is that they laugh at prof, in my opinion it seen rude to other people.

My cultural collection:

In our country, there's fewer classes like the class in this story because students have less communications with teachers. Mostly, the teachers will stand in front of us and talk for whole 40 minutes.

My cultural collection:

this story told try to make the #1 jar full, But, why me should full my life? I just want to live in a peace. and relax place. I don't want the weight on my back.

... what do you do if you find difficult words or sentences?

• 请2-4位老师分为一组，每人拿出一张白纸，确定该题目的评分标准：

1) 学生的作答当中可以看到哪些核心素养及具体要素？

(construct)

2) 每个维度总分设置，及给分点如何确定？

3) 试着给以下两个同学的回答进行评分。

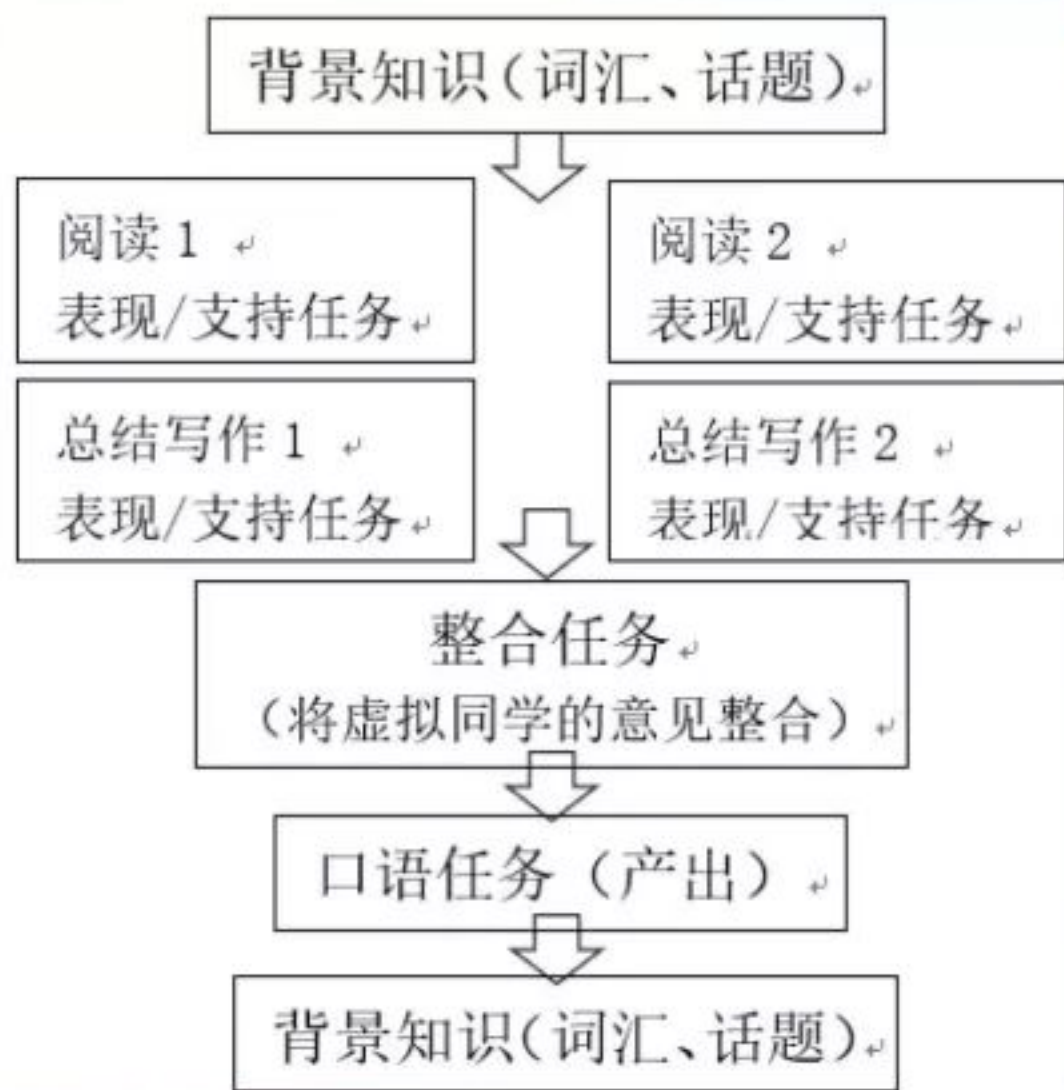
## (2) SBA: 基于情景的评价

SBA是ETS发起的基于认知的评价学习、促学以及为学（Cognitively Based Assessment of, for, and as Learning简称CBAL）研究项目框架中所提出的评价方法。

SBA利用科技创设真实的语言使用情境，设计一系列有序的任务和试题来检测使用者完成某项任务的能力，并在过程中提供一些脚手架支持，从而探测使用者在完成任任务过程中的学习成长(O'Reilly & Sabatini, 2013; 2019)。

“This assessment (SBA) thus provides a **perfect opportunity** for assessing the display and development of L2 proficiency, topical knowledge, and reasoning skills in which contextual factors, L2 resources, topical resources, socio-cognitive, and dispositional resources convene to play an explicit role in task achievement.”  
(Purpura, 2017, p. 57)

One technology-based innovation that would lend itself to integrated assessments of language and content are scenario-based assessments. ... Both content and language scaffolds could be incorporated into scenario-based assessment. (Llosa, 2017: p. 11)



SBA任务模型举例Purpura (2018)

## The Group Project

Your class has decided to enter a contest for a free trip abroad. Your teacher has allowed you to work with three friends to research some options and create a **travel proposal**.

You will send the final proposal to the organizers of the contest, and they will select the best proposal.

The illustration shows two scenes. On the left, a group of people is sitting around a table in a meeting room, labeled 'Your group'. An arrow labeled 'Send Travel Proposal' points to the right, where a group of people is sitting around a table in a meeting room, labeled 'Organizing Committee'. A small portrait of a man is visible in the bottom right corner of the illustration area.

You




Mr. Perez

What do you already know?

1. Which country was established by the Nabataeans in the 1<sup>st</sup> century B.C?
- A) Jordan
  - B) Vietnam
  - C) I don't know.

2. Where can you find the Bedouin community?
- A) Jordan
  - B) Vietnam
  - C) I don't know.

3. Which picture shows people eating Manna?
- A) 

- B) 

**PERFORMANCE  
MODERATOR:**  
Background  
Knowledge

**Purpose:**  
Measure prior  
background  
know  
(topic)



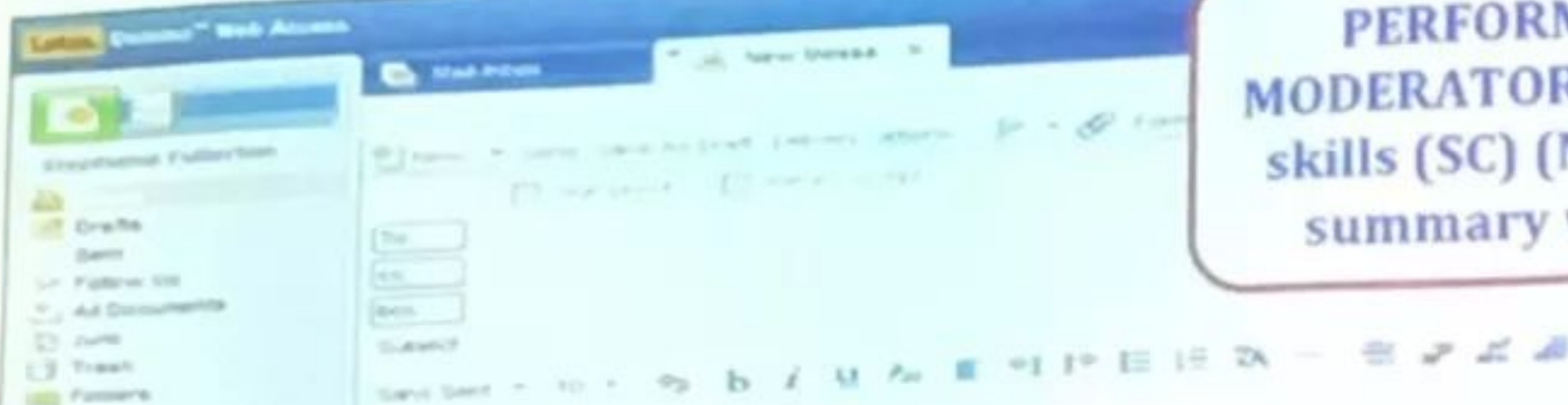


Anna

Hi everyone, I already summarized the cons in an email. Can you summarize pros and send it to me? I'll put them together and send the whole thing to Mr. Perez? Take a look at mine before you write yours, OK? We should have the same style.



Mr. Perez



**PERFORMANCE MODERATOR:** Enabling skills (SC) (Modelling summary writing)

Anna's summary of the cons goes here as a model for writing a summary



Reading

Purpose Reading implied n & pragma

A



## SBA是否能够有效服务于英语评价目标？

语言教学及评价目标

阅读功能及阅读评价

基于情景的阅读评价

- SBA由ETS研究者提出特征：
- 利用网络创设真实情景
- 话题内容知识、认知能力
- 在做任务过程中，逐渐提供支持，强调以评促学

## 重点二：真实性

核心素养强调关键能力和品格，而这些在真实情境中解决具体问题时才能够更好地体现，因此，需要强调评价的**真实性**。（程晓堂，2018）

Authentic test tasks are in the sense that the testees must be able to recognize the communicative purpose of a task in order to respond appropriately.  
(罗少茜, 2009:225)



# • 哪个更真实？ 哪个更能测评学生的能力素养？

0 You can eat here in the mornings.

Answer:

0	A	B	C	D	E	F	G	H
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

1 You should not swim here.

2 You must not drive fast here.

3 You can play football here after lessons.

4 It is cheaper to buy things today than tomorrow.

5 You can drive here next week.

A

**SLOW!  
DANGEROUS CROSSROADS**

B

SWIMMING POOL  
OPEN AFTERNOONS  
Adults - £2.50  
Children - £1.00

C

HALF PRICE FOOTBALL SHIRTS -  
SALE MUST END THIS AFTERNOON

D

**POLICE CARS ONLY**

E

**DANGER!  
DO NOT GO INTO  
THE WATER**

F

BREAKFAST SERVED  
7.00 - 10.00

G

**ROAD CLOSED  
UNTIL WEEKEND**

H

SCHOOL SPORTS CLUB  
NOW OPEN IN THE  
EVENINGS!



A



B



C

1. \_\_\_\_\_ a kind of things you wear to protect your eyes from bright sunlight ✓

2. \_\_\_\_\_ an object similar to an umbrella you carry to protect yourself from sunlight ✓

3. \_\_\_\_\_ a pair of light shoes you wear when you walk on the beach ✓

• To think:

• 相对真实的任务具有哪些特征？这些特征为何可以更好地探测学生的能力？

## 哪种方式更符合真实问题情境？哪种方式更能体现出学生的素养？哪些素养？

- 假设你是李华，你想找一个笔友（pen pal）。请根据下列提示写一篇不少于100字的作文。

(1) 自我介绍；

(2) 你为什么想找一位笔友？

(3) 如何与你联系？

青少年应该从小就学会辨别是非。有人认为，为了帮助小孩子辨别是非，惩罚是必要的。↵

你同意以上观点吗？你认为应该允许家长和教师使用哪些惩罚手段来教育孩子分清是非？↵

请你以普通读者为对象，就以上话题写一篇短文，回答以上问题并给出理由。↵

## 总结：变与不变之**变**（I）

- 未来教育测量将会迎来以下变化：
- 测量我们认为重要的**能力**（例如，增加社会-情绪学习）；
- 为收集这些能力的证据而设计测评任务（按照更丰富的学习模型编制更复杂的任务；采用更多的背景观察）；
- 把收集到的证据与考生的特征联系起来（例如，用新的学科模型）；
- 通过更好、更具有互动性的报告在决策中交流测量结果；
- 评价测量的质量和影响，例如，更大程度地关注测量对教和学的正面作用，以及公众对测量的积极反映。

## 变与不变之**不变**(2)

- 测量的基本特性**不太可能发生变化**：**四个方面**
  - **Assess what we aim to assess: Competence.** 进行观察，收集证据：包括测量设计、项目编制和施测。
  - 测量模型：将观察到的证据与考生个体、群体或者组织机构有意义的特征联系起来（并估计与这些特征联系的不确定性）。
  - 在决策或者分数报告方面进行交流，应用测量的结果。
  - 评价测量的效度：评价证据收集的机会、特征以及决策的质量和影响。

(Bennett, 2018)

## 参考文献

- Bennett, Randy E. (2018). Educational Assessment: What to Watch in a Rapidly Changing World. *Educational Measurement: Issues and Practice, Winter, Vol. 37, No. 4*, pp. 7–15.
- 程晓堂. (2017). 英语学科核心素养及其测评. 中国考试, (5): 7-14.
- 程晓堂. (2018). 基于问题情境的英语考试命题理念与技术. 《外语测试与教学》
- 程晓堂, 赵思奇. (2016). 英语学科核心素养的实质内涵. 课程·教材·教法, (5): 79-86.
- 教育部 (2018). 《高中英语课程标准》。北京: 人民教育出版社。
- 罗少茜、黄剑、马晓蕾. (2014). 促进学习: 二语教学中的形成性评价. 北京: 外语教学与研究出版社.
- 罗少茜、张玉美. (2018). 运用阅读圈任务评价学生英语学科核心素养. 中小学外语教学, (7), 1-5.
- 罗少茜、张玉美、赵海永 (出版中)。《英语测试与评价》。北京: 外语教学与研究出版社。
- 梅德明、王蔷. (2018). 改什么? 如何教? 怎样考? 高中英语新课标解析. 北京: 外语教学与研究出版社.





Thank you!