

Guoxing Yu, PhD  
Reader in Language Education & Assessment

# Integrated language assessment tasks

Nottingham, 13 April 2018



---

---

---

---

---

---

---

---

2

45:00

## Outline of the presentation

- What is integrated language assessment: premise, promises, problems and compromises? What do “integrated language assessment” tasks look like in English language tests – using summarization (central to successful performance in any integrated writing/speaking task) as an example
- What are the major trends/foci of existing research on integrated language assessment?
- Call for collective effort and holistic approach to exploring integrated assessment systematically: A list of questions and potential topics for study



---

---

---

---

---

---

---

---

3

45:00

## Influence of major English language tests

- Current interest and research efforts on integrated assessment have been sparked to a large extent by the introduction of integrated assessment tasks (writing and speaking) in TOEFL iBT; however,
- Integrated assessment is not a recent invention
- The old IELTS in 1990s, and many other tests



---

---

---

---

---

---

---

---

### Defining integrated assessment tasks (1)

- *Tasks that require learners or test takers to incorporate substantive content from source materials in print, audio, and/or visual forms (Cumming 2013, p.1)*




---

---

---

---

---

---

---

---

### Defining integrated assessment tasks (2): terminologies

- Writing from sources, source-based writing
- Content-responsible writing
- Reading-to-write, listening-to-write
- Discourse synthesis
- Summary writing




---

---

---

---

---

---

---

---

### Defining integrated assessment tasks (3) - Integrated or integrative test?

- **A pretty old term:** integrative test (Dictionary of Language Testing, 1999) – a test in which learners are required to **combine various skills** in answering test items as opposed to a discrete-point test in which each item focuses on a single element of language




---

---

---

---

---

---

---

---

### Defining integrated-ness from different perspectives

- **SKILL:** Integration of different language skills (e.g., listening-to-speaking, listening-reading-writing)
- **CONTENT:** Integration of language and content
- **SOURCE:** Integration of sources of multi-modalities (e.g., video, texts, graphs, see Cog-Pro1 and 2 and 3, funded by British Council on IELTS AWT1, & GEPT in Taiwan funded by LTTC)




---

---

---

---

---

---

---

---

### Five promises of integrated writing tasks (Cumming 2013)

- Provide realistic/authentic, challenging literacy activities
- Engage test takers in writing that is responsible to specific content
- Counter test method or practice effects associated with conventional item types
- Evaluate language abilities consistent with construction-integration or multi-literacies models of literacy
- Offer diagnostic value for instruction or self-assessment




---

---

---

---

---

---

---

---

**Summarization is central to successful performance in integrated tasks**




---

---

---

---

---

---

---

---

### Why integrated assessment tasks (summary-writing as an example): Premise

- Reading comprehension is a *sine qua non* for summarization (Yu 2005);
- Summarization is a natural entailment of reading comprehension (Kintsch & van Dijk 1978, van Dijk & Kintsch 1983), though not necessarily automatically (Johns 1983)
- *Summarization requires the comprehension, evaluation, condensation, and frequent transformation of ideas that have been presented.*  
(Hidi & Anderson 1986: 473-74)




---

---

---

---

---

---

---

---

### Promises (1)

- "the **hub** in the wheel of reading comprehension" (Axelrod 1975: 383)
- "an **essential** communicative activity" (Brown and Smiley 1978)
- Summarization tasks have a **natural appeal** because they "simulate real-world tasks in which non-native readers have to read and write a summary of the main ideas of a text" (Cohen 1994: 174)
- Summarization tasks: "a **very common** exercise" in pedagogy (Seidlhofer 1995:2), **popular** in reading textbooks (see Weir et al. 2000)




---

---

---

---

---

---

---

---

### Promises (2)

- Summarization: the **very act** of university students' life (Allison et al 1994, 1995a, b; Friend 2001)
- Summarization skills essential for **content acquisition and academic success** (Friend 2001, Maclellan 1997, Holmes & Ramos 1993, Rea-Dickins, Kiely & Yu 2007)
- "Whatever a person's interest in studying a foreign language, there seems to be **no escape** from the acquisition and development of summarising skills"  
(Johns 1988:79)




---

---

---

---

---

---

---

---

### Promises (3)

- “a test requiring test takers to...summarize the **propositional content** in a reading passage, will involve the **full** [emphasis added] range of organizational characteristics” (Bachman 1990: 139)




---

---

---

---

---

---

---

---

### The five perils of integrated writing tasks

- Confound the measurement of writing abilities with abilities to comprehend source materials
- Muddle assessment and diagnostic information together
- **Involve genres that are ill-defined and so difficult to score** (what exactly is summary writing, for example?)




---

---

---

---

---

---

---

---

### The five perils of integrated writing tasks

- Require threshold levels of abilities for competent performance, producing results for examinees that may not compare neatly across different ability levels;
- Elicit texts in which the language from source materials is hard to distinguish from examinees' own language production (productive vs re-productive, Yu 2013)




---

---

---

---

---

---

---

---

**Perils of integrated assessment (using summary-writing as an example), Problems in task directions:**

- *"Summarize the points made in the lecture you just heard, explaining how they cast doubt on the points made in the reading.... Your response will be judged on the basis of the quality of your writing and on how well your response presents the points in the lecture and their relationship to the reading passage" [TOEFL iBT]*




---

---

---

---

---

---

---

---

---

---

**Read the passage below and summarize it using one sentence. Type your response in the box at the bottom of the screen. You have 10 minutes to finish this task. Your response will be judged on the quality of your writing and on how well your response presents the key points in the passage.**

'Just-in-Time' is a method of manufacturing products which aims to minimise production time, production costs, and the amount of stock held in the factory. Raw materials and supplies arrive at the factory as they are required, and consequently there is very little stock sitting idle at any one time. Each stage of the production process finishes just before the next stage is due to commence and therefore the lead-time is significantly reduced. With a 'just-in-time' production system, the level of production is related to the demand for the output (i.e. the number of orders) rather than simply producing finished goods and waiting for orders. This means that raw materials and stock only need to be ordered from suppliers as required.

Just in time is...]

Cut

Copy

Paste

Total Word Count: 4

Next (N)

---

---

---

---

---

---

---

---

---

---

**You will hear a short lecture. Write a summary for a fellow student who was not present at the lecture. You should write 50-70 words.**

Status: Playing

Volume

Cut

Copy

Paste

Total Word Count: 0

Next (N)

---

---

---

---

---

---

---

---

---

---

19 **You will hear a recording. Click on the paragraph that best relates to the recording.**

Status: Playing

Volume

Recent research shows that when customers are dissatisfied with service or products, they usually inform the company of their complaint. They tend to only talk about their negative experience with others if the company doesn't respond to their initial complaint.

Recent research shows that nearly one third of dissatisfied customers will share a negative retail experience with someone else. Nearly half of the people surveyed said that they had avoided a store due to another individual's negative experience.

Recent research shows that nearly 80% of customers who have a bad experience with a company share that experience with people they know. Furthermore, 70% of people who hear about another individual's negative experience will avoid a store based on what they've heard.

Next

---

---

---

---

---

---

---

---

---

---

---

---

20 **Summarise the information by selecting and reporting the main features, and make comparisons where relevant. Write at least 150 words.** 45:00

Mode of Transport	1960	1980	2000
bus	22	28	18
car	8	25	38
bike	28	24	8
foot	36	18	10

CAER University of BRISTOL

---

---

---

---

---

---

---

---

---

---

---

---

21 **国际人才英语考试** 45:00

高级(书面沟通), Task 2为撰写报告。  
考生阅读一份商务图表, 按要求撰写一份150词左右的报告。本任务考查考生描述、比较、概括图表关键信息的能力。- see also IELTS Academic

高端- 口头沟通

**Task 1**

Read the following passage and then listen to a 2-minute lecture on the same topic from a different perspective. Summarize the points made in the lecture, and explain how the points respond to those made in the reading passage. Note-taking is allowed while you read and listen.

You will have 2 minutes to read the passage, 2 minutes to prepare and 90 seconds to speak.

CAER University of BRISTOL

---

---

---

---

---

---

---

---

---

---

---

---

### What is a summary?

- “A serious problem in much of this literature, however, is the assumption that only one kind of summary exists” (Ratteray 1985: 457).
- **sequential summaries** that retain the original order in which information was presented (including *abstract, précis, secretarial minutes, abridged digest*) and
- **synthesizing summaries** that alter this sequence to achieve specific objectives (including *locational digest, restructuring digest* and *review*).




---

---

---

---

---

---

---

---

### Transparency in task directions and assessment criteria (1)

- Clear and transparent task directions are essential to ensure that test takers know exactly **what kind of summary [integrated tasks]** they are expected to produce and what to include and not to include in the summary. We can't just tell test takers to “**summarize it**”. (or **use the source texts**)
- It's imperative for item writers and teachers to take into account the various factors that can contribute to summarization performance. **[integrated tasks]**




---

---

---

---

---

---

---

---

### Transparency in task directions and assessment criteria (2)

- Summarization **[integrated writing tasks]** may well be a unique construct, it's important to employ parameters or indicators different from and additional to those for independent writings in order to better measure the quality of a summary in both automatic evaluation systems and marking schemes for human raters.
- Differences in terms of “**productive language**” between independent and integrated writings (= *re-productive*) – a major concern (copy-paste vs **creativity** in writing)




---

---

---

---

---

---

---

---



### Current focuses/trends of research on integrated language assessment tasks (1)

- Compare the differences in students' performance between independent and integrated tasks
- the processes or strategies used by test takers during the integrated tasks (writing and speaking)
- Analyse the discourse features of written or spoken texts produced in literal and critical summaries of source materials (listening and/or reading), e.g., use of source texts, intertextuality,
- Role of reading and writing proficiency in integrated




---

---

---

---

---

---

---

---

### Current focuses/trends of research on integrated language assessment tasks (2)

- Instructors or raters' perceptions of integrated tasks
- Integrated assessment for diagnostic values
- Comparing L1 and L2 integrated tasks (Zhu, Li, Yu, et al. 2016; and Zhu & Yu research in progress funded by UGC HK), task performance, and cognitive processing




---

---

---

---

---

---

---

---

### Call for collective effort and holistic approach: a series of questions

- **Conceptualisation of integrated assessment:** What can be counted as "integrated assessment" tasks? What does integrated-ness mean? With what reference point do we define integrated-ness?
- **INPUT SOURCES:** What are the differential impacts of the features of the source input (e.g., visuals, audios, texts, graphs, paper-based or computer-delivered) on task performance?




---

---

---

---

---

---

---

---

**FILTER PLANT:**

- What are the cognitive processes involved in completing integrated assessment tasks?
- What roles do test takers' characteristics (e.g., language and scientific skills, social, educational background, training experience, and interest in the source texts) play in their performance?




---

---

---

---

---

---

---

---

**EVALUATION**

- What would be the key indicators of success/failure of integrated task performance?
- To what extent should the evaluation criteria differ from those for independent tasks?
- How to design data-driven rating scale for integrated reading-to-write scale? (Erwert & Shin 2015)




---

---

---

---

---

---

---

---

- What are rater reactions to integrated tasks and rating scales, how do raters make decisions (Gebriel & Plakans 2014)
- How to develop rubrics (Chan et al. 2015)




---

---

---

---

---

---

---

---

- How do we operationalize the multidimensionality of integrated tasks from different perspectives, for different assessment purposes, users, test takers, and etc.?
- What are the practicality issues of implementing integrated assessment in large-scale tests, at design, administration and evaluation stages?




---

---

---

---

---

---

---

---

**TEACHING, COACHING/LEARNING AND IMPACT**

- To what extent are integrated tasks coachable?
- Where is the demarcation between desirable and unethical intensive coaching for the test tasks?
- What are the impacts of integrated assessment on language education more generally?
- What are students' practices and abilities for writing from sources (Cumming et al. 2018)




---

---

---

---

---

---

---

---

- To what extent has **CEFR** (the scales of writing and speaking proficiency) reflected/incorporated "integrated assessment"?
- To what extent has **China Standards of English** (the scales of writing and speaking proficiency) reflected/incorporated "integrated assessment"?
- To what extent can **TECHNOLOGY** help us better understand integrated assessment tasks?
- What are the synergies between EAP writing research and LANGUAGE TESTING research?




---

---

---

---

---

---

---

---

Guoxing Yu, PhD  
Reader in Language Education & Assessment

Email: [Guoxing.Yu@bristol.ac.uk](mailto:Guoxing.Yu@bristol.ac.uk)

Skype: i-TESTS



---

---

---

---

---

---

---

---