Guoxing Yu, PhD
Reader in Language Education & Assessment

Integrated
anguage
assessment tasks

Nottingham, 13 April 2018

University of
BRISTOL

45:00

### 2

## 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/focuses 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





45:00

### 3

# 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





45:008 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) University of BRISTOL **CÆR** 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 **CÆR** University of BRISTOL 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 discretepoint test in which each item focuses on a single element of language

University of BRISTOL

**CÆR** 

# 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 multimodalities (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)





45:008

# 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





45:00

# Summarization is central to successful performance in integrated tasks





45:008 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) University of BRISTOL 45:008 Promises (1) "the hub in the wheel of reading comprehension" (Axelrod 1975: "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) **CÆ**R University of BRISTOL 45:00 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"

University of BRISTOL

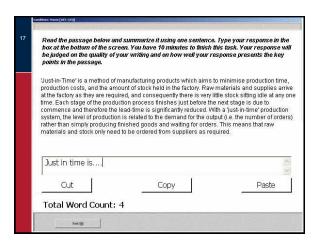
(Johns 1988:79)

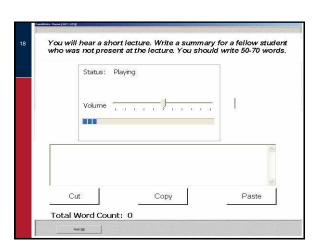
45:008 Promises (3) • "a test requiring test takers to...summarize the propositional content in a reading passage, will involve the full [emphasis added] range of organizational characteristics" (Bachman 1990: 139) University of BRISTOL **CÆR** 45:008 The five perils of integrated writing tasks · Confound 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?) **CÆR** University of BRISTOL 45:00 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 reproductive, Yu 2013)

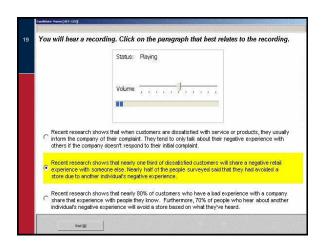
University of BRISTOL

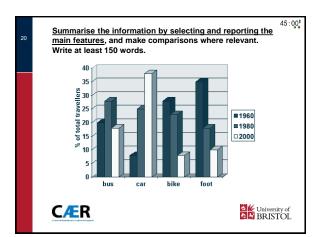
**CÆ**R

# Perils of integrated assessment (using summarywriting 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]











	45:0 <mark>0</mark> \$	
22	What is a summary?	
22	<ul> <li>"A serious problem in much of this literature, however, is the assumption that only one kind of summary exists" (Ratteray 1985: 457).</li> <li>sequential summaries that retain the original order in which information was presented (including abstract, précis, secretarial minutes, abridged digest) and</li> <li>synthesizing summaries that alter this sequence to achieve specific objectives (including locational digest, restructuring digest and review).</li> </ul>	
23	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]	
	Summanzation penormance. [integrated tasks]	
	CÆR University of BRISTOL	
24	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.	
	<ul> <li>Differences in terms of "productive language"</li> <li>between independent and integrated writings (= re-</li> </ul>	
	productive) – a major concern (copy-paste vs	
	Geat R in writing)	

25	Current focuses/trends of research on integrated language assessment tasks (1)	
	Compare the differences in students' performance	
	between independent and integrated tasks	
	<ul> <li>the processes or strategies used by test takers</li> </ul>	
	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	
	<ul><li>of source texts, intertextuality,</li><li>Role of reading and writing proficiency in integrated</li></ul>	
	Ctask Rerformance	
		7
00	Current focuses/trends of research on	
26	integrated language assessment tasks (2)	
	<ul> <li>Instructors or raters' perceptions of integrated tasks</li> </ul>	
	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	
	<b></b>	
	CÆR University of BRISTOL	
		<u> </u>
	45:00 <sup>8</sup>	1
27	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?	
	C = D	
	CÆR University of BRISTOL	

45:008 • 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? **CÆR** University of BRISTOL 45:008 **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) **CÆR** University of BRISTOL 45:00 · What are rater reactions to integrated tasks and rating scales, how do raters make decisions (Gebril & Plakans 2014) · How to develop rubrics (Chan et al. 2015) **CÆR** University of BRISTOL

45:00 · 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? **CÆR** University of BRISTOL 45:00 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) **CÆ**R University of BRISTOL 45:00° · 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? **CÆR** University of BRISTOL

Guoxing Yu, PhD Reader in Language Education & Assessment	
Email: <b>Guoxing.Yu@bristol.ac.uk</b> Skype: <b>i-TESTS</b>	
University of BRISTOL	