

# The IELTS Speaking Test: what can we learn from examiner voices?

---

Chihiro Inoue, Nahal Khabbazzashi,  
Daniel Lam & Fumiyo Nakatsuhara

*CRELLA, University of Bedfordshire*

# Acknowledgements

This presentation draws upon a research project funded by the British Council as part of the Joint-funded Research Programme. Any opinions, findings, conclusions or recommendations expressed in this material are those of the presenters and do not necessarily reflect the views of the British Council or its partners.

Special thanks to:

- Mina Patel
- Steve Copeland
- Barry O'Sullivan and
- Over 1,200 IELTS Speaking examiners who participated in this study

# Research Background

# History of IELTS Speaking Test (IST)

**EPTB:** English Proficiency Test Battery

**ELTS:** English Language Testing Service

**IELTS:** International English Language Testing System

**IELTS:** After the Speaking Test Revision Project (1998-2001)

1965-81

1981-89

1989-2001

2001-present

# English Proficiency Test Battery (EPTB;1965-81)

- Reading & Listening comprehension test; **No speaking component**

# English Language Testing Service (ELTS;1981-89)

- 10-15 mins **f2f interview**
- **Subject specific:** 5 domains (Life Sciences, Social Studies, Physical Sciences, Technology, Medicine) + General Academic
- **3 parts:** (1) interview (2) subject-specific discussion  
(3) discussion of future plans
- **Integrated test** of reading into speaking

## IELTS (Original design – 1989-2001)

- 10-15 mins **f2f interview**
- **No link to specific domains:** *Measuring 'general proficiency in speaking'*
- **5 parts:** (1) introduction (2) extended discourse (3) elicitation (4) speculation and attitudes (5) conclusion
- **No rigid interlocutor frame**
- **Holistic rating scale**

## IELTS (After the 2001 Revision)

- 11-15 mins **f2f interview**
- Assessing **general proficiency in speaking**
- **3 parts:** (1) interview (2) long turn (3) discussion
- Examiner training programme with the use of the **interlocutor frame**
- **Analytic rating scales** with 4 categories

**Over 15 years since the last revision**



**WHAT NOW?**

**IELTS**<sup>TM</sup>  
English for International Opportunity

# IELTS and Examiner voices

- **ELTS Revision Project (1986-89):** Stakeholder questionnaires & interviews → no major advantage due to too varied target stakeholders + unfocussed data collection (Davies, 2008)
- **Merrylees & McDowell (1999):** Examiner survey on a wide range of aspects (N=151) → contribution to the IELTS Speaking Revision Project (1998-2001)
- **Brown & Taylor (2006):** Examiner survey on a wide range of aspects (N=269)
- **Brown (2006):** Examiner interviews on rater perceptions and rating process (N=6)
- **Galaczi, Lim & Khabbazzashi (2012):** Examiner survey on rating experience and perception of the rating scales (N=1142)
- **Nakatsuhara, Inoue, Berry & Galaczi (2017a, b):** Examiner interviews, focus groups on test delivery and rating aspects (N=14)

## Lessons:

**Well-defined survey items, focussed data collection, Quan+Qual, Clear link with the literature**



# Aims of the Project(s)

## THIS STUDY

To gather IELTS Speaking examiners and examiner trainers' voices on various aspects of the current IST and what changes they would like to see.

To  
sy  
lit  
sp  
ge  
rec  
the

## TOWARDS NEW AVENUES FOR THE IELTS SPEAKING TEST: INSIGHTS FROM A COMPREHENSIVE LITERATURE REVIEW

Fumiyo Nakatsuhara, Chihiro Inoue, Daniel Lam & Nahal Khabzbashi  
Centre for Research in English Language Learning and Assessment,  
University of Bedfordshire, UK

4 June 2018

To offer a range of possibilities and recommendations for the next revision of the IELTS Speaking Test.

# Research Questions

**RQ1:** What are the IELTS examiners' and examiner trainers' views towards the IELTS Speaking Test and their suggestions for future improvement?

**RQ2:** What similarities and differences can be discerned between the recommendations based on the literature review (from the other project) and the results of the examiner survey and interviews (RQ1)?

# Methodology

# Sequential Mixed Methods Design

Stage 1

**Focus group**  
with 3 examiners +  
previous research  
+ Input from the  
IELTS Partners

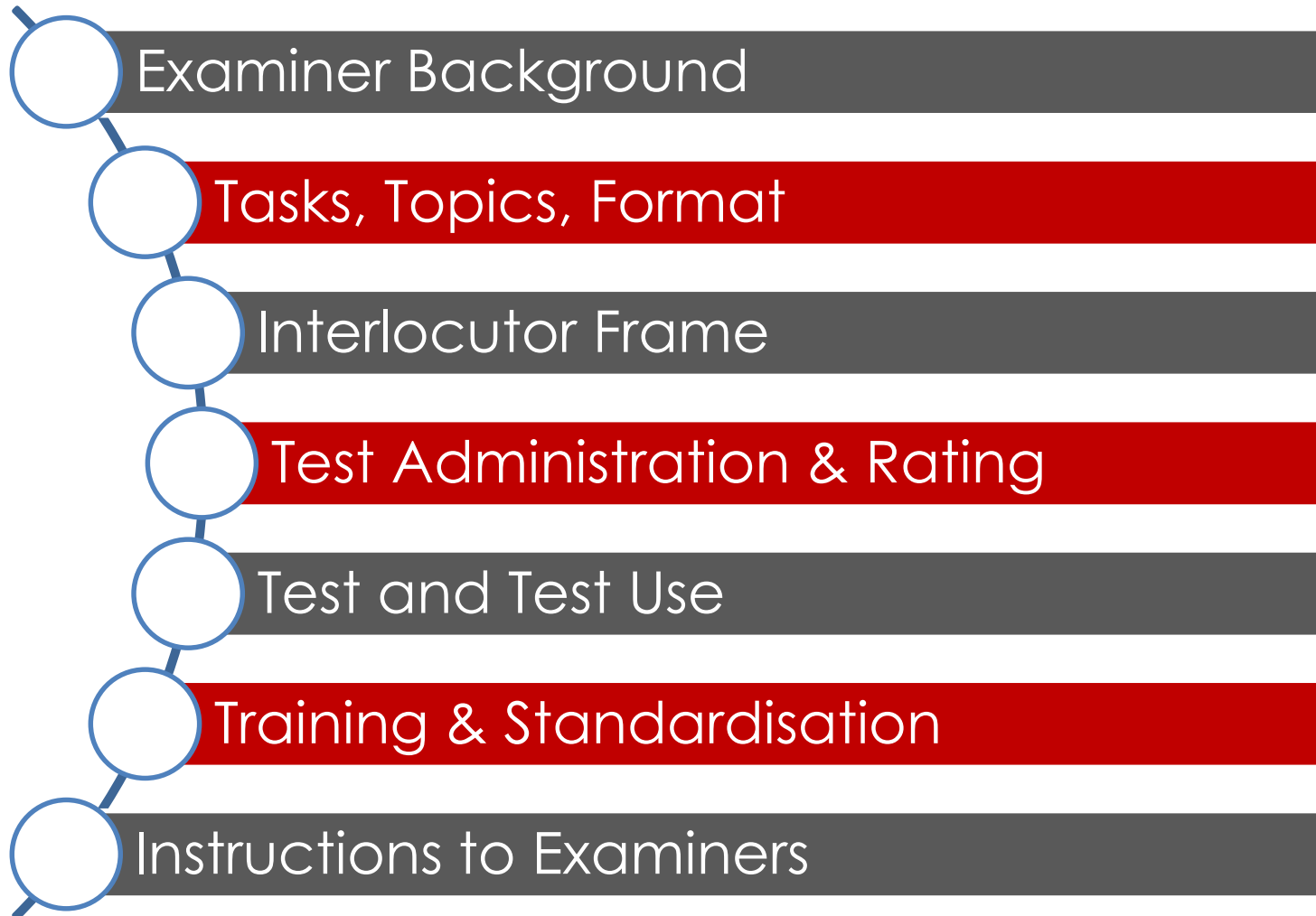
Stage 2

**Online survey**  
with 1,203  
examiners all over  
the world

Stage 3

**Semi-structured  
interviews**

# Online Survey



# Sequential Mixed Methods Design

Stage 1

**Focus group**  
with 3 examiners +  
previous research  
+ Input from the  
IELTS Partners

Stage 2

**Online survey**  
with over 1200  
examiners all over  
the world

Stage 3

**Semi-structured  
interviews**

# Participant Selection

- Call for volunteers at the end of the survey
  - 418 volunteered to be interviewed
  - Sampling on the basis of region, examining experience, and diversity of expressed opinions
- *interview data representative of examiners' voices*
- Interviews with **30** examiners & **5** trainers
    - *Novice to highly experienced (less than 6 months to more than 23 years)*





# Semi-structured Interviews

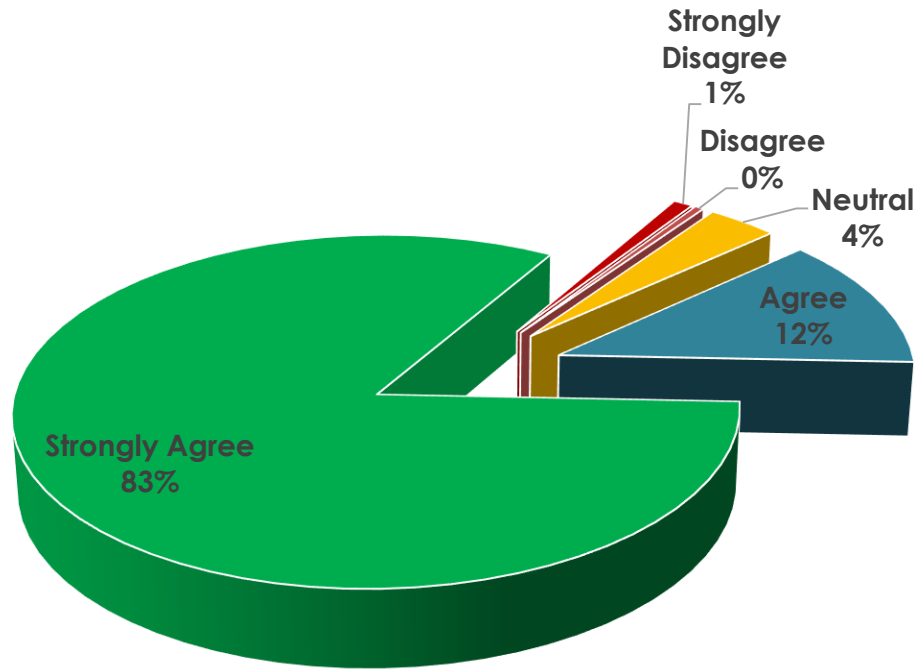
- Based on survey responses
- Tailored to individual examiners
- Covered different areas of interest
- Video/audio calls  $\cong$  1 hour

**→ *Thematic analysis of transcripts***

# Results (1)

**Face-to-face vs Computer mode**

# Face-to-Face vs. Computer



**Survey results:** 95% of examiners agreed that the f2f interaction mode is more suitable for test delivery compared to a CB mode.

# Authenticity and construct representation



'We have an interview because we are interested in communicative abilities and skills that you cannot get from other things. It's like you are cutting your nose to spite your face; in essence you have an interview because you can't test in a computer.'



'Computers can't replace human interactions. Gestures, eye contact, etc. are all parts of language ability. The purpose of the speaking test is to test candidates' ability to speak in a natural communicative environment.'



- Acknowledgement of artificial nature of all assessments
- Time limitations, speaking to a computer → part of construct
- **BUT** the 'human' element seen as strength of the exam; a degree of 'naturalness' while CB assessment is 'one more step removed from what language is about'

## Scepticism surrounding CB solutions



'Until technology is good enough a human has to be in charge of it. Otherwise you'll be messing around with the kids.'

'Computer-based? Even BBC that is probably using the best technology gets words wrong. A scenario where the software has difficulty with exact words by British native speaker on the news, how do you expect our guys from Pakistan? From the Philippines? Or our friends from Scotland? Case closed!'

## Providing support to candidates

'When f2f with another person you have lots of options to support a candidate whether it is facial gesture like a smile or a hand to say 'continue' but the computer does not do that.'

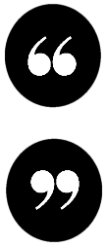
## Reducing cheating

'I used to work in China where there was a TOEFL test with a computer speaking component and they are good at working out what the questions are; they used to prepare, and memorise and the kind of answers was completely rote so those same students in real life...their speaking skills were horrible and they just memorised. And you can challenge them better with the f2f test.'



## Reducing stress and test-related anxiety

'I see a lot of pitfalls and lots of stress with the speaking part of the TOEFL – They are worried about so many things and having to talk into the computer...and there's the timing issue that IELTS doesn't have. And that's a good thing for candidates.'



'We have to remember that most people are very nervous and a human voice can be very reassuring and having someone face to face can be really helpful.'

'I have taught TOEFL preparation and they are very different. TOEFL does not give leeway for emotional reactions, or being sick running out of the room but a f2f interaction makes the student much more relaxed. With IELTS you can skip questions or take your time and go as slow and fast as you like. F2f in general is much more calming in general but computer-based can be very jarring.'

# Video-conferencing technology as a possible alternative

'Testing remotely, seeing your face on the screen, I guess is a viable option and second best thing but having an anonymous without authentic interaction and a computer voice, we lose a lot. A test that tests human interaction is a marker of what we need to do in the real world.'

'It's not like I dislike technology but what you've got to realise that we are not numbers. We are individuals and we want a **human element.**'



**Future direction of the IELTS Speaking Test**



# Results (2)

## Interlocutor frame

# Survey

# Interviews

Part 1 – a bit too rigid / too rigid (62.1%)

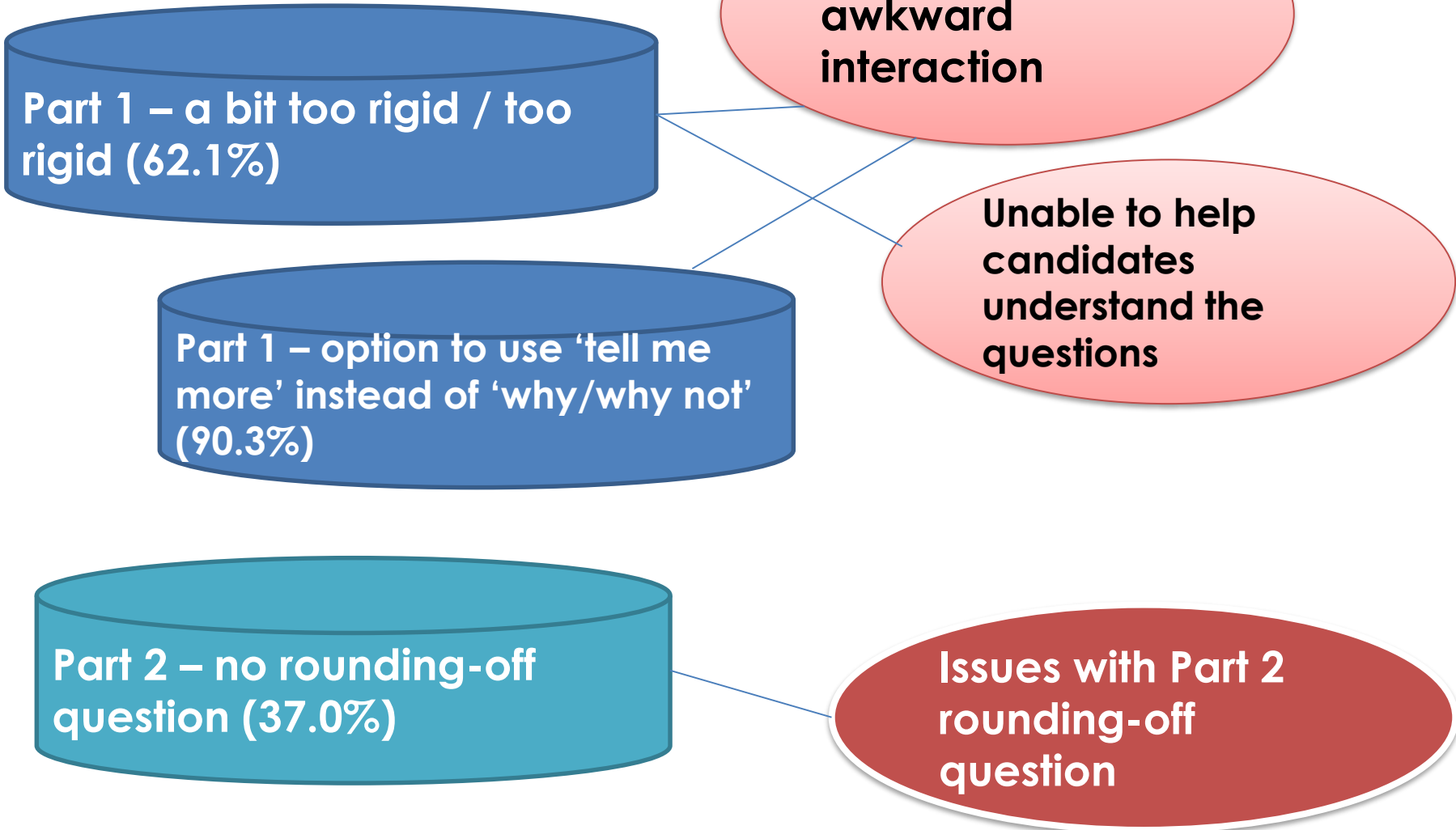
Part 1 – option to use 'tell me more' instead of 'why/why not' (90.3%)

Part 2 – no rounding-off question (37.0%)

Inappropriate/  
awkward  
interaction

Unable to help  
candidates  
understand the  
questions

Issues with Part 2  
rounding-off  
question



# Inappropriate/awkward interaction

C: Sorry, in my religion, music is not allowed.  
Int: What type of music do you like?

C: I have studied 6 years of medical school.  
Int: Why?

Why/  
why not?

Int: Do you always carry an umbrella when it's raining?  
C: Yes.  
Int: Why?

Some candidates pre-empt the 'why' question by including a short reason/explanation.

If candidate gives short, 1-sentence answers, after 5-6 of those, it starts sounding like interrogation, and it's intimidating.

# Unable to help candidate understand the question

[In Part 1] If they say I don't understand the question, examiners **can only repeat the question**. The only thing examiners are allowed to do [in Parts 2 & 3] is to give a short gloss of the word **if the candidate asks**.

Sometimes Part 3 Qs are too difficult, I don't understand what they want..... **You want to** give them an example, a suggestion, **but you're not supposed to**.

As ET, I have come across some less strong examiners who can't or don't explain /paraphrase the vocabulary – the candidate [ended up] being **entirely silent...**

# Part 2 Rounding-off Questions

- Why examiners would rather not have them
  - Already ‘answered’
  - Irrelevant
  - Not meant to elicit ratable language sample
  - Pressure of time limit

You ask a question, the candidate thinks that they're in an English test, so when being asked a question, they try and give a long, detailed answer, and then the **examiner had to cut them off** after 10 seconds because they have run out of time in Part 2.

# Dilemma

Not asking  
the rounding-  
off question?



Going over  
time?

**FAIL**

# Conclusions

# Human interaction as a key feature of IELTS Speaking Test

- Examiners are generally in favour of F2F mode of speaking test for its construct representation
- A rigid interlocutor frame, while a well-intended and appropriate response to issues in the previous test version, strips away the element of 'human interaction' that the IELTS Speaking Test can take pride in
- Integrating video-conferencing technology is a possible alternative
- Judicious flexibility with interlocutor frame in the spirit of the test is welcomed



“**Balance** the need to **standardise** the test event as much as possible ...against the need to give examiners **some degree of flexibility** so that they ... feel that the language of the event is **natural** and free flowing”

(O’Sullivan and Lu, 2006: 22)

**Thank you!**