

Assessment of Candidates' Interactional Competence Using Group Oral Tests

Fumiyo Nakatsuhara
CRELLA, University of Bedfordshire, UK

Research Background

Desirability of the group oral tests

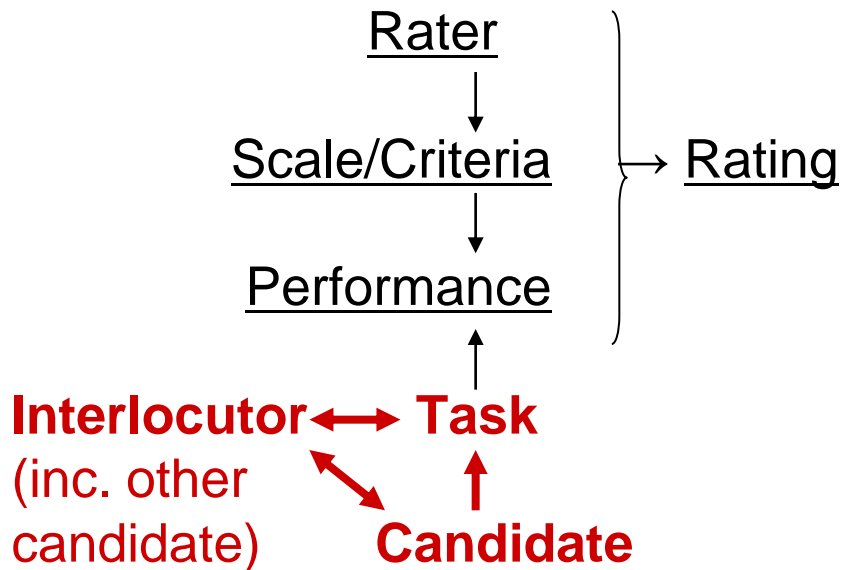
- 1) Quick and efficient way of testing spoken language (Bonk & Ockey, 2003)
- 2) Less burden on examiners (Folland & Robertson, 1976); more consistent test-administration (Ockey, 2001)
- 3) Symmetrical contribution to the interaction (Van Lier, 1989); eliciting richer language samples (French, 2003);
- 4) Students' positive perception (Fulcher, 1996; Ockey, 2001); possible positive washback effect (Hilsdon, 1995)

Current practice of group oral tests

- The Hong Kong AS Level Examination
- The College English Test, China
- The Kanda English Proficiency Test (KEPT), Japan

Impact of test-taker characteristics

Co-constructed nature of speaking test performance
(McNamara, 1996)



- Gender
- Acquaintanceship
- Cultural background
- L1
- Personality
- Proficiency level

(e.g. O'Sullivan, 2000; Berry, 2004; Norton, 2005; Ockey, 2006; Van Moere & Bonk, 2004)

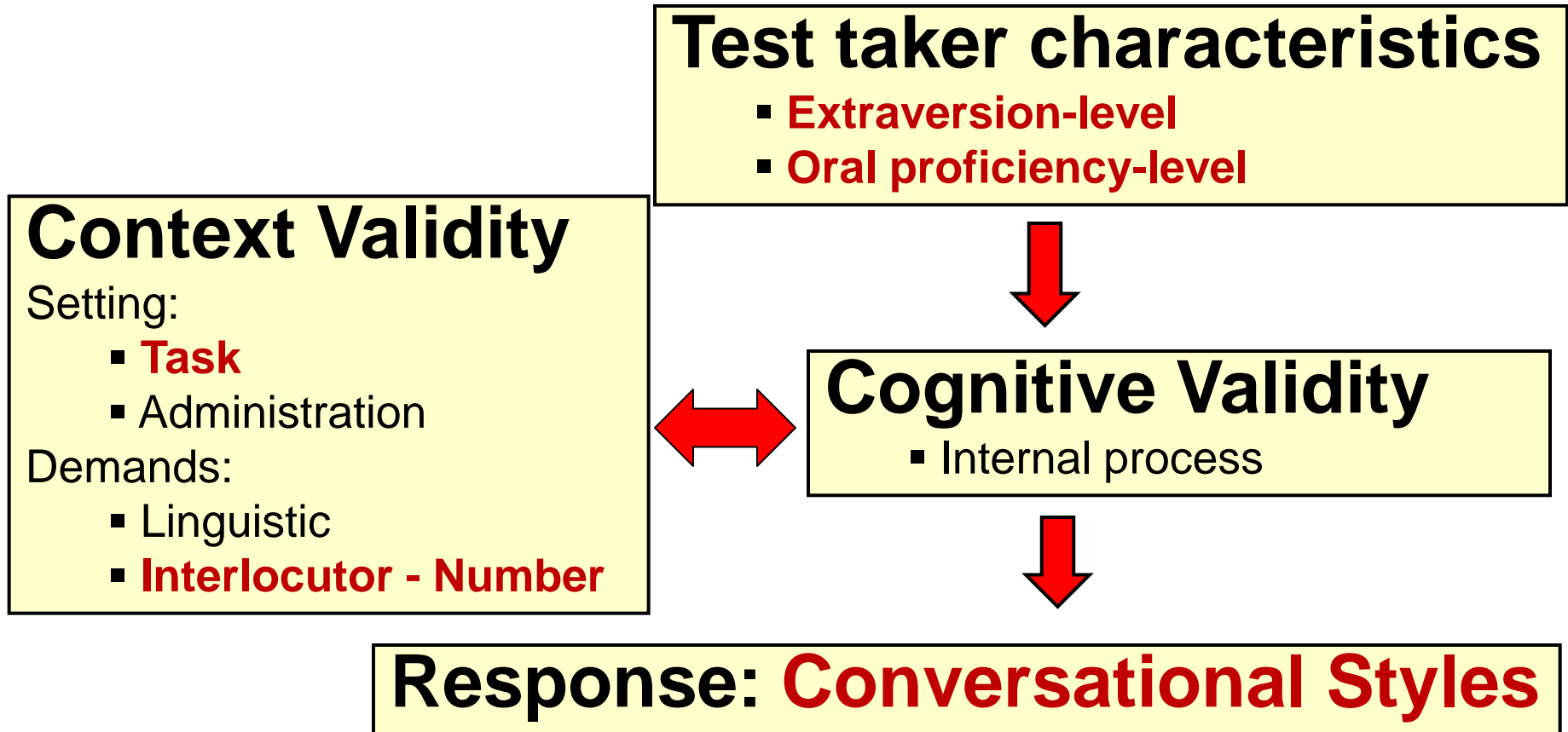
- Results are often mixed in terms of the direction of the effects

Some more issues to be investigated

- **Impact of test-taker characteristics on discourse**
 - Most studies on group tests have examined scores (e.g. Van Moere & Bonk, 2004; Ockey, 2006), and how people talk in group tests is unclear
- **Impact of task types**
 - Most studies have employed a free discussion format, and very little is known about other group testing tasks.
- **Impact of group sizes**
 - Grouping students into groups of 3 is better than groups of 4 [G4 → 1 person tends to be left out] (Ojima, 2005).

Socio-cognitive framework for validating speaking tests

Weir (2005)



Research Questions

RQ1: Impact of two test-taker characteristics in general

- Are conversational styles in groups affected by a test-taker's own and his/her group members' **extraversion- and proficiency-levels**?

RQ2: Impact of two test-taker characteristics & task types

- Do test-takers' extraversion- and oral proficiency-levels have different influences on conversational styles **among different task types**?

RQ3: Impact of two test-taker characteristics & group sizes

- Do test-takers' extraversion- and oral proficiency-levels have different influences on conversational styles in **groups-of-three** participants as against **groups-of-four**?
- ❖ If there are any influences/differences, **how & why** do they occur?

Methods of Data Collection & Data Analysis

Data Collection

- **Participants:** 269 Japanese high school students
- **Grouping:**
 - Grouping students into **groups of 3 or 4** as they wish (controlling for **acquaintanceship** and **gender**)
- **Test-taker characteristics:**
 - **Extraversion-level:** a Japanese version of Eysenck Personality Questionnaire (EPQ) (Iwawaki *et al.*, 1980)
 - **Oral proficiency-level:** classroom teacher's assessment
- **Tasks:**
 1. **Information-gap task** more closed/more structured
 2. **Ranking task** ↕
 3. **Free discussion task** more open/less structured

Your class has **20,000 yen** to spend by **next March**. It was agreed to use this money to buy a **camera**. Your classmates asked your group to **decide what camera to buy**. You have information about *Camera C* and *D*, and your group members have information about *Camera A*, *B*, *E* and *F*.

(1) Exchange all the information you have and (2) decide which camera your class is going to buy. (Note: you can use all 20,000 yen, but you may want to save some money for other things.)

When you have finished discussing, please tell me **which camera you would buy and why you want to buy it.**

Camera A
Camera F



Camera B



Camera C



Camera D



Camera E



	Price (yen)	Weight (gram)	Type of camera	Other characteristics
Camera A				
Camera B				
Camera C	16,000	91	Digital camera	You can take movies. (2 Mega pixels)
Camera D	600	90	Disposable camera	You can take 27 photos with flash.
Camera E				
Camera F				

It is said that a good high school teacher should have the following qualities.



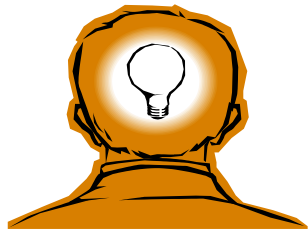
Knowledge of subject



Clear speaking voice



Ability to organize class



Intelligence



Pleasant appearance



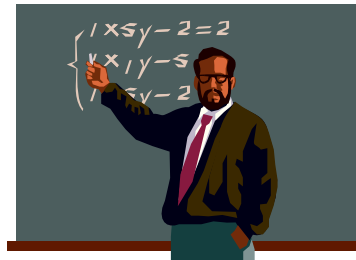
Enthusiasm for teaching



Sense of humor



Fairness



Clear writing



Love of students

(1) Discuss **how important** these qualities are.

(2) Decide **which three qualities are the most important**, and give the **reasons** .

When you have agreed on the **three qualities**, please tell me **which three things are the most important** and **why you thought so**.

Your friend, Ken, will have the first date with his girlfriend tomorrow.

He came to your group and asked a question.

Should boys pay all the costs for dates?

Here are some opinions from other friends.



Boys and girls are equal. Neither girls nor boys have the right to ask the other to pay the bill.

Many girls like boys who are generous. But, if I keep paying for all the dates, I will be poor and I won't be able to see my girlfriend any more.



Some boys like paying for dates. If they pay, they can usually plan the date and impress girls.

Generally speaking, boys eat a lot!



Please **discuss your opinions** so that you can **give him suggestions** on what he should do tomorrow.

When you have finished discussion, please tell me what you, as a group, would like to suggest him.

Data Analysis

Quantitative Analysis (Multiple Regression)

■ Predictors (IVs)

- **Extraversion-level: Japanese EPQ (0-20)**
 1. Self (own) score
 2. Self-excluded group mean
 3. Self-excluded group Std.Dev.
- **Oral proficiency-level: classroom teacher's assessment (0-5)**
 4. Self (own) score
 5. Self-excluded group mean
 6. Self-excluded group Std.Dev.

■ Measures of Conversational Styles (DVs)

- **Goal-Orientation:** measured by Topic initiation
- **Interactional Contingency:** measured by Topic ratification
- **Quantitative Dominance:** measured by The amount of talk

(Van Lier, 1989; Young & Milanovic, 1992; Young, 1995; Kormos, 1999)

Qualitative Analysis (Conversation Analysis)

To interpret and elaborate on the quantitative results

Quantitative & Qualitative Results

- **MR: Model Summary [Overall]**
- The given model (with six IVs) accounted for **14%** of the variance in **topic initiation** and **21%** of the variance in **the amount of talk**, but *not* topic ratification.

DV	R Square	Sig.
Topic initiation	.144	.000
Topic ratification	.004	.955
The amount of talk	.208	.000

MR: Main Finding 1 [Overall]

- A more extraverted/more proficient test-taker **initiated more topics and talked more**, especially when they were grouped with less extraverted/less proficient group members.

How & why?

❖ Scaffolding behaviour in expert/novice asymmetric interactions

[Excerpt 1] 4008 (E:15, P:4) 4024 (E:0, P:2) 4032 (E:16, P:3)

1 → 4008: So, let's discuss about qualities ((looking at 4024 and 4032)). (.5) First one,
2 knowledge of subjects? (.5) What do you think about it? ((looking at 4032))

3 4032: Uh, Huh huh [uh

4 4008: [Wha(h)t?

5 (.8)

6 4032: Uh: uh::: uh: uh I think knowledge of subject is very important.

7 4008: Yeh ((nodding)), of course.

8 4032: Uh huh

9 (1.0)

10 → 4008: How about clear speaking voice? ((gesturing towards 4024))

11 4024: I think this is more important (.5) for us to clear speaking voice.

12 4008: uh huh ((nodding))

❖ Use of body language

(e.g. putting a hand towards to the next speaker, raising a hand, clapping hands)

- (a) to involve quiet members (b) to solve interactional problems (c) to draw an attention

❖ Tolerance of silence

[Excerpt 2] 3013 (E:8, P:3) 3017 (E:14, P:3) 3028 (E:7, P:3)

1 3017: Uh: I organize, ability to organize class is, a- what mean ability to
2 organize class?

3 (6.0)

4 →3017: Wha(h)t does mean ability to organize class?

5 (4.0)

6 →3017: Organize mean *soshiki*? *Soshiki*?

7 3028: ((nodding))

8 (3.0)

9 →3017: I (.5) I e? most important is love of student, but it is

MR: Main Finding 2 [Task-type comparison]

- **The proficiency-level variables** were influential in all tasks.
- **The extraversion-level variables** were more influential **in more open tasks** (the ranking and the free discussion tasks), but *not* in the info-gap task.

How & why?

Info-gap Task

- ❖ **Compulsory info exchange → Making all test-takers talk**
- ❖ **Information order forcing the interactional order & role**

[Excerpt 3] 5001 (E:8, P:3), 5006 (E:14, P:3), 5007 (E:16, P:3)

5 5001: I know about camera A. Uh It's price is uh

:

15 5007: I have camera C. This camera is sixteen thousand yen.

:

21 5006: Camera E is two: two thousand ye- uh?

:

25 (3.0) ((5006 & 5007 looking at the prompt card, while occasionally throwing a glance at 5001))

26 5001: uh

27 (2.0)

28 5006: huh huh

29→5001: What do you like uh? What do you uh:: would you like , what what

30 would you like to, Midori?

Ranking & Free Discussion Tasks

❖ Personalising the given topics

(a) to justify opinions (b) to persuade other group members (c) to involve quiet speakers into the personalised stories

Free Discussion Task

The discourse agenda was not formulated →

- ❖ The liveliness of the interaction differed greatly (e.g. involvement of **jokes**)
- ❖ **Difficult to use well-functioning sequence openers** (in Expert/Novice asymmetrical interaction)

[**Excerpt 4**] [1071 \(E:16, P:2\)](#) [1077 \(E:18, P:3\)](#) 1081 (E:6, P:1) 1083 (E:4, P:0)

1 → 1071: Do you think? What do you think? [**←unspecific opener**]

2 1083: Huh huh huh huh

3 (1.5)

4 → 1077: What do you think? [**←unspecific opener**]

5 (3.0)

6 → 1071: If you you:: (.5) you go to restaurant, restaurant with a boy, (.5) ah you
7 pay or money pay

8 1081: Pay [me?

9 → 1071: [pay or boy pay? Which do you like? [**←specific opener**]

10 1081: Half and half.

MR: Main Finding 3

[Group-size comparison]

- **The proficiency-level variables** were influential in both group sizes.
- **The extraversion-level variables** were more influential in **groups of 4** than in groups of 3.

How & why?

❖ Collaborative atmosphere in groups of 3 → Mitigating the effect of extraversion variables

- Joint utterance completion in groups of 3

- Avoidance behaviour in groups of 4

[Excerpt 5] Group of 4, 5045 (E:0, P:3) 5046 (E:16, P:4) 5047 (E:14, P:4) 5049 (E:1, P:3)

1 5046: What do you think? ((making deliberate eye contact with 5049))

2 → 5049: Me too.

[Excerpt 6] Group of 4, 3002 (E:3, P:3) 3022 (E:5, P:1) 3026 (E:12, P:) 3032 (E:12, P:5)

1 3032: Do you have any any () anything else?

2 (8.0)

3 3026: huh huh *Ryoko*?

4 → 3022: Do you think about clear speaking voice, *Azumi*?

In addition...

❖ The turn-taking was sometimes mechanical in groups of 4

[Excerpt 7] Specifying turn-taking order by gesture in groups of 4

1 2104: Have you ever been (.) have you ever going to date, date?

2 (1.0)

3→ 2106: ((indicating to take turns in a counter-clockwise way))

[Excerpt 8] Irrelevant use of “How about you?” in Groups of 4

1→ 5002: How about you, Maya?

2 5001: Ah:: I I think clear writing is uh is important, because uh my teacher

:

6→ 5001: S(h)o I think clear writing is important. **How about you, Midori?**

7 5007: Uh:: I think love of students is important, because when ah teacher

:

10→5007: How about you, Yukari?

Incompatibility between talking naturally in groups of 4 and talking in groups of 4 in oral tests

- The test interaction does not allow “*schisming*” (Egbert, 1997)
- → unconsciously avoiding the simultaneous talk
- → inducing the unnatural way of turn-taking

Conclusion

- **A test-taker's characteristics, his/her group members' characteristics, task types, group sizes** affected the resulting test-takers' discourse in group oral tests.
 - the **interactionalist view** of construct definition (e.g. Brown, 2005)
 - Greater attention should be paid to **task types & group size**

- **[in general]** A more extraverted/more proficient test-taker **initiated more topics and talked more**, especially when they were grouped with less extraverted/less proficient group members.

- **[task] The extraversion-level variables** were systematically more influential in more open tasks.

- If extraversion-level variables are not within the test construct, **the info-gap task** could be the most preferable...
 - However, if group oral tests are **to test communication ability**, extraversion-level should be part of the test construct (considering the reasons for the impact).

- **[group size]** Interactions in **groups of 3** seem more suitable for group oral tests.
 - Groups of 4 could make it more difficult to elicit ratable speech from introverted test-takers.

- Incompatibility between talking naturally in groups of 4 and talking in groups of 4 in oral tests
 - **Grouping test-takers into groups of 4 might not always provide a suitable environment where test-takers could display their communication ability!!**

- To refine **‘interactive communication’ rating scales** for the co-constructed interaction
 - *Not ‘shared scoring’ (e.g., May, 2007)*
 - Rating scales that take the dynamics of the interaction process into account (Chalhoub-Deville & Deville, 2005: 826)

- **Brown (2003: 20)**

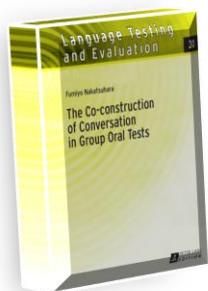
“It is...simply not appropriate to assume that the variation [of performance] that is allowed to occur is not relevant to the construct, especially where the construct can be interpreted as encompassing interpersonal communication skills.”

- **Swain (in Fox, 2004: 241-242)**

“I think we make the assumption that if you put two people together in the same testing setting, that they would automatically feel competitive toward each other, when in fact ... they helped, supported and scaffolded each other.”

Thank you very much!

Fumiyo Nakatsuhara
University of Bedfordshire
Fumiyo.Nakatsuhara@beds.ac.uk



Nakatsuhara, F. (2011) Effects of the number of participants on group oral test performance, *Language testing* 28 (4): 483-508

Nakatsuhara, F. (2013) The co-construction of conversation in group oral tests, Peter Lang.