

【研究論文】

**A Handbook for the BECC Bunkyo English Speaking Test -
Preparation and Procedures**

**広島文教大学BECC Bunkyo English Speaking Test
面接官のためのハンドブック-準備と手順**

広島文教大学 文教英語コミュニケーションセンター

講師 **Richard Sugg**

Abstract

The Bunkyo English Communication Center (BECC) Bunkyo English Speaking Tests (BESTs) and their associated teacher training and standardization materials were developed between 2015 and 2017. However, the updating of the Cambridge A2 Key English Test and its next level B1 Preliminary English Test on which they were based forced us to look again at the kind of paired-speaking assessment used within our General English (GE) curriculums. The resulting decision to reorganize the make-up of the Sophomore English (SE) twice-yearly end of semester speaking tests, and to break up the equivalent Freshman English (FE) tests into smaller speaking assessments unfortunately coincided with the outbreak of the 2019 Novel Coronavirus (COVID-19). The resulting cancellation of all paired assessments led to a loss in overall institutional knowledge and a disparity in teachers' speaking test experience. The following handbook is a first step to rectifying this problem as we resume the SE end of semester tests and look forward to finally implementing the proposed 'new for 2020' in-course paired speaking assessment from the spring of 2023.

Introduction

The development of the Bunkyo English Speaking Tests (BESTs) and their associated teacher training and standardization between 2015 and 2017 has already been covered within this university's publications (Sugg & Svien, 2018). However, the updating of the Cambridge A2 Key English Test (Cambridge, 2020a) and its next level B1 Preliminary English Test (Cambridge, 2022b) forced us to look again at the kind of paired-speaking assessment we wanted to use with our General English (GE) curriculums. Based on the Cambridge changes we decided to reorganize the makeup of the Sophomore English twice yearly end of semester speaking tests, and to break up the equivalent Freshman English (FE) tests into smaller speaking assessments based on the SE speaking tests' individual parts. However, the outbreak of the 2019 Novel Coronavirus (COVID-19) resulted in all paired-speaking assessments, student preparation, and staff training sessions being put on hold until the summer of 2022. This two-year delay resulted in a loss of overall institutional knowledge, and with the inevitable changeover in staff members over that time led to a disparity in teachers' speaking test experience. While the online materials and standardization page, and the resumption of the in-house training session in July 2022 helped us successfully complete this academic year's first semester SE tests, it was impossible to review all previously acquired knowledge at the same time

as updating everybody with the new changes. Consequently, and with the second semester SE tests in mind, the need for a regularly updated 'how-to' guide became apparent.

This handbook, while still a work in progress, is designed to be used by all Bunkyo English Communication Center (BECC) staff before administering the twice-yearly Bunkyo English Speaking Tests (BESTs) and is to be referred to during the assessments as required. It is also intended to be a basis for reference for future General English Course in-class paired speaking assessments.

Contents

1. General Information
 2. Assessment Level
 3. Format and Breakdown
 4. Timing
 5. Workflow
 6. Assessing the BEST
 7. Extra Advice and Tips for Interlocutors and Raters
 8. Room and Desk Set-up
 9. Students with Special Requirements
 10. Absent, Late, or an Odd Number of students
 11. Test Security
 12. Interlocutor and Rater Checklists
 13. Appendices
- Appendix 1: Interlocutor Rubric
- Appendix 2: Holistic Rubric
- Appendix 3: Example Student Roster and Rater Grades Sheet.
- Appendix 4: CEFR ratings: What is A1, A2, or B1?
- Appendix 5: Interlocutor Script and Questions
- Appendix 6: Student Mark Sheets

1. General Information

The BESTs are the Bunkyo English Speaking Tests. These used to be held at the end of the first and second semesters of the Freshman and Sophomore years. As of 2022 ONLY the Sophomore students will take the BESTs. The Freshman students will take in-class short speaking assessments based on BEST Part 1 or Part 2 at specified times during their first and second semesters.

The BESTs are designed to fulfill the following three main purposes:

- To find out if our students are achieving the GE curriculum speaking proficiency goals.
- To provide us with an objective measure of speaking ability.
- To provide us with a speaking assessment aligned to the CEFR.

In turn, the results of the BESTs will be used to assign fair grades to students, give feedback to students on their speaking ability, and to provide feedback on the curriculum to the Curriculum Coordinator. To ensure that the tests are both objective and aligned to the Common European Framework of Reference (CEFR), the current versions referred to in this document have been based on a modified version of the updated 2020 Cambridge Key English Test (KET) (Cambridge, 2020a). This test, as with all Cambridge Assessment Cambridge Suite tests, is supported by validated research and has many freely available reference and training materials. As per these exams, the BESTs are double rated by an interlocutor and a rater, and all grades are analyzed with the Rasch Facets program, with the goal of providing fair scores for all students. (Sugg & Svien, 2018)

2. Assessment Level

The Cambridge Key English Test (KET) is an examination at the pre-intermediate level on the Cambridge Assessment English Overall Speaking Scale. At the CEFR A2 level (Cambridge, 2020a), students who fulfill the requirements have a basic command of the spoken language and can cope with a range of everyday situations. However, a score of 5 on the KET speaking test is equivalent to a basic pass level on the Cambridge Preliminary English Test (PET). At this intermediate CEFR B1 level (Cambridge, 2020b), students have a limited but effective command of the spoken language and can handle communication in most familiar situations. In the same way, the BESTs aim to differentiate students at the CEFR A1 level (cannot fulfill the KET requirements), A2, and B1 levels. Further clarification of the characteristics of students at these levels can be found in Appendix 4.

3. Format and Breakdown

The BEST format is as follows:

- Students take the test in pairs.
- Students are rated by two people (interlocutor and rater).

There are two, two-part tasks, adapted from the updated 2020 Cambridge Key English Test to be completed **within 10 minutes**.

The BEST breakdown is as follows:

Part 1

CEFR Components: Overall oral interaction and understanding an interlocutor. Students answer simple personal questions, extended questions on areas of personal interest, while taking part in a basic information exchange with an interlocutor.

CEFR A2. Students can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Students can understand enough to manage simple, routine exchanges without undue effort (COE, 2018).

- **Up to 4 minutes.** Students talk with the interlocutor.
- **Phase 1** – up to 1.5 minutes. Students answer simple ‘warm up’ questions *What’s your name? Whose English class are you in? etc.*
- **Phase 2** - Three questions. two are short answer (hopefully a sentence) and the longer third question is to get three or four utterances about a topic of personal interest.
- **Back up questions are only used if a student cannot answer.** If they can only give one sentence answers to the third question, that is already a clue to their level.
- **ALL the question options for this part have been seen by the students on the BEST practice lesson.** This is to ensure that even Pre-A1 and A1 students can do the majority of Phase 1 and Phase 2 short questions and can therefore participate in the assessment.

Part 2

CEFR Component: Spoken interaction, overall oral production via a conversation. Students are asked to express likes and dislikes, while at the same time giving simple reasons and interacting with a partner. The conversation is centered around a card with visual prompts but does not have to come to a conclusion.

CEFR A2. Students can participate in short conversations in routine contexts on topics of interest (COE, 2018).

CEFR B1. Can enter unprepared into conversation of familiar topics, express personal opinions and exchange information on topics that are familiar, of personal interest or pertinent to everyday life (COE, 2018).

- **Up to 6 minutes.** Students talk with each other THEN the interlocutor.
- **Phase 1** – Students have up to 2 minutes to discuss the things on the card. This is from THEIR perspective. What are the things, do they like them? Can they say why or why not and engage in general discussion about that topic.
- **This is followed by some very basic adjective questions.** These are asked to both students at once. Answers are short: yes and why, or no and why not. The student that doesn’t answer is then asked, ‘what do you think, (*name*)?’ Students should answer at least one question each, but more if time permits. If nobody volunteers an answer, then direct one question to each student.
- **Round up questions:** there is a round-up question that is asked to each student.
- **Phase 2** - 2 questions. These are longer questions asked alternately to both candidates. The idea is to get three or four utterances related to the Phase 1 topic.
- **You can repeat the questions to the second student IF NECESSARY. No reforming or back up questions.** If a student can’t answer, just say ‘that’s ok, thank you.’
- **NONE of the question options for this part have been seen by the students in the BEST practice lesson.** Pre-A1 and A1 students may well start to struggle on Part 2. Shorter answers for them are ok. Superior performance on the card and on the Phase 2 questions will mark out the B1 students.

4. Timing

Each BEST pair interview is to be completed within 10 minutes. It is important for reasons of test reliability and fairness that all tests are of the same length, and that each component is of the same length.

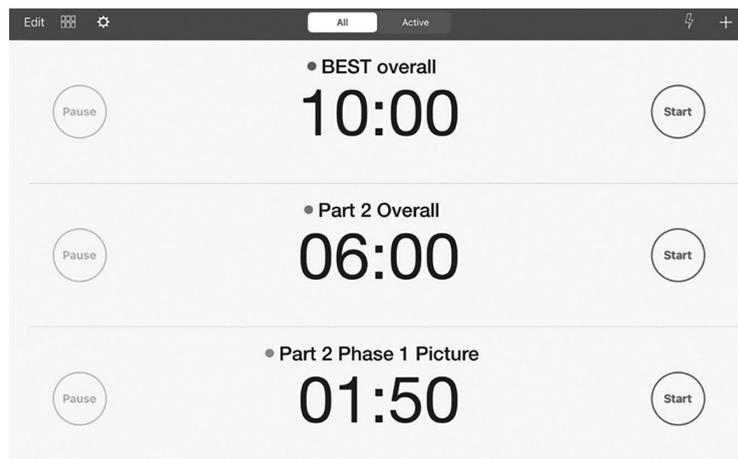
- The ten minutes start as the pair sit down. **Part 1 Phase 1** usually takes **within 1.5 minutes**.
- **Part 1 Phase 1 and Phase 2** should take **no longer than 4 minutes**.
- **Part 2 Phase 1** students have up to **2 minutes** to try and discuss the card.
- **Part 2 Phase 2** should be completed **within** the 10 minutes.

To help interlocutors keep to these timings, they should use a silent timing device such as an iPad with an app capable of tracking multiple concurrent timings. Our current app of choice is Timer+.

This should be set up with three timers:

- **Timer 1** BEST Overall 10 minutes.
- **Timer 2** Part 2 Overall 6 minutes.
- **Timer 3** Part 2 Phase 1 Picture 1minute 50 seconds.

Here is an example of a correctly set up timer:



Please refer to the BEST Workflow for an overview of the interview timings.

5. Workflow

Rater (Analytic Rubric)	Interlocutor (Holistic Rubric)	🕒
Calls the next pair of candidates Reminds candidates to speak up Directs candidates to the chairs		
Goes to the rater seat	Ask candidates to sit START 10-minute timer	0:00
	Part 1 Phase 1	
# IF required, starts the video camera.	Asks candidates for their mark sheets	
During Part 1: Form a grade hypothesis	Introduces themselves and the rater.	
	Follow the script: ask for name, student number, from, department, and whose class.	
	Part 1 Phase 2	1:00 -1:30
	Follow the script: Ask candidate B the two short answer questions (substitute question IF necessary). Ask the longer answer extended question (shorter hint questions if necessary). Repeat for Candidate A.	
During Part 2: Test grade Hypothesis / check for sustained performance	Part 2 Phase 1 START 6-minute timer	3:00 -4:00
	Follow the script: Introduce Part 2 Phase 1. Place the card in front of the students.	
	START 1 minute 50 seconds timer	
	Sit back to encourage students to talk to each other	
Acknowledge Interlocutor Sample OK / a greater sample required	Check with Rater that they are happy with the language sample obtained from Part 2 Phase 1	
	Follow the script: Ask adjective questions as appropriate.	
	Follow the script: Ask round up questions.	
	TAKE BACK THE CARD	
	Part 2 Phase 2 (up to 2 minutes, within the total 10)	6:00 -8:00
	Follow the script: Ask the first question to candidate B. Redirect to A (Only repeat the question if necessary)	
Complete grade assessment	Follow the script: Ask the first question to candidate A. Redirect to B (Only repeat the question if necessary)	
Escort candidates back and encourage them to leave.	Thank the candidates and send them away.	~10:00
Call the next pair when the interlocutor is ready.	Check student mark sheets are completed and put out of sight. Prepare for next pair.	

6. Assessing the BEST

The BESTs are double rated by an interlocutor and a rater. The interlocutor uses the holistic rubric (Appendix 1), and the rater uses the analytic rubric (Appendix 2). These rubrics are based on the CEFR and are amalgams of freely available rubrics from Cambridge (2020a, 2020b), EIKEN, TEAP and IELTS. In assessing the students, we are trying to focus on what students CAN DO, rather than looking to take points off for students making mistakes. The holistic rubric consists of one set of descriptors and provides a score out of five points. The analytic rubric consists of three sets of descriptors and provides a score out of fifteen points.

A key thing to keep in mind when assessing is the three-step performance process: hypothesis, test, and sustained performance.

1) Hypothesis

As you listen to Part 1 Phase 1 and Phase 2 short questions, form your idea of the test taker's level. (Start at 4, then look above / below)

2) Test

During Phase 2 extended questions and Part 2 Phase 1 (card), focus on the descriptors for that band level. Test your hypothesis. Does the student meet, exceed, or fail to meet those descriptors.

3) Sustained performance

Confirm (or deny!) that the level of performance is maintained in the Part 2 Adjective, round up, and Phase 2 questions. **It is vital** to test your initial hypothesis, and check for sustained performance, because students often start strongly on the memorized part of the test, but then tail off when extra content or explanation is required. However, **be careful of** other factors which may cause the 'opposite' effect, where students start quietly and then produce more language later on. These include nervousness, being paired with a sibling (training video 2), someone from another course, someone from the same course with whom they don't get on, or someone they consider to be 'senior' to themselves.

When using the **holistic rubric**, some points to remember are:

- It is better to start looking at a '4' level, and then move up or down depending on a student's aptitude. Remember the key is '**sustained performance**'.
- To get a 2, a student would not be able to answer most of **the part 1** questions, as well as **not** be able to do part 2.
- A student **does not have to be fluent** to get a 5. Read the rubric and concentrate on the **bolded** areas.

Generally, if a student **CAN DO**

- Part 1 Phase 1 and Phase 2 short, then they are at least A1(3).
- Part 2 Phase 1 using sentences then they are at least A2 (4).

However

- In Part 2, students need longer utterances to be B1 or above.
- Students need to answer the Part 1 Phase 2 extended questions and Part 2 round up and Phase 2 questions well to be B1 or above.

Well can still involve using simple language. Please remember to follow the descriptors in the rubric

When using the **analytic rubric**, some points to remember are:

- It is better to start looking at a '4' level, and then move up or down depending on a student's

aptitude. Remember the key is 'sustained performance'.

- To get a 2, a student would **not** even be able to answer most of **the part 1** questions, as well as not be able to do part 2.
- Gestures: Is the student using gestures **naturally to go with** their words (5), OR is a student using them **in lieu of words** (4), OR is a student **relying on gestures** (3)?
- A student **does not have to be fluent** to get a 5. However, students should be able to perform Part 2 **well** to get a 5 for Grammar and Vocabulary (GV) and for Interactive Communication (IC).

Some extra guidance for each set of descriptors is as follows:

Grammar and Vocabulary (GV)

- Band 2 = student uses mainly Japanese or is unintelligible.
- Vocabulary **range** = three or more related words.

Pronunciation (P)

- If you can't hear a student, then they can be considered a 2 for pronunciation.

However, please remember to remind every student before their interview to speak up, and to position yourself as close as is acceptably possible. You may move your desk forward once during an interview if you still feel it is necessary.

Interactive communication (IC)

- Ability to maintain exchanges is the **main** consideration.
- Count the responses if a student seems weak early on.
- **Interlocutors** provide **prompting**.
- **Partners** provide **support**.

The interactive communication grade will not necessarily be the same as the interlocutor's grade, as the interlocutor and rater are working independently and are grading in a different way.

7. Extra Advice and Tips for Interlocutors and Raters

Interlocutors

When speaking to students, go for a slower, smoother, 'rhythmical' style. Sudden changes in pace or volume can be very off-putting. Please remember to look at the correct candidate when you are speaking to them, and to lean back from the table at the beginning of Part 2 Phase 1. This helps to indicate that the students should interact with each other, rather than with you. You can also gesture when you say 'now speak together' to re-enforce this.

For Part 1 Phase 2 (short) questions, only ask the 'back up' for that question **if** students say nothing at all or ask for repetition. If a student gives at least a 1 sentence answer, move on.

For Part 1 Phase 2 (extended) questions, ask the question and listen to the answer. If two of the 'points' covered in the script are covered, move on. If a student goes silent in this part, asks for repetition, or only answers one point, you can ask **up to** all three prompt questions depending on time. Try to balance getting a ratable sample against time and causing stress to students.

Please **stick to the script**, as this maintains validity and fairness for each pair. Do not pick the script up and have it between yourself and the students while speaking. Look down at the script, then **look up** to say the lines. It is better to be familiar with the script, rather than to have it memorized. Even after five or six pairs, please don't try to speak from memory as this invariably leads to differences

in question formation or even mistakes.

With practice and preparation, mistakes can be kept to a minimum. However, if they do happen, don't panic. If you catch the mistake right away – mispronouncing a student's name or stumbling over a question – apologize, repeat correctly, and move on. If you realize after a short period of time – 30 seconds or more – unless the students have reacted to the mistake, or have been confused by it, continue with the interview as it is. If you feel it has affected the students, then apologize and go back to the point where the mistake occurred.

It is good to have your rubric memorized, but you should still refer to it during the interviews. Our online training, in-house training, and this manual are all aiming to standardize teachers. There will of course be some variation in grading, but the Rasch Facets analysis can mitigate these differences if they are consistent.

Raters

One of the jobs of the rater is to keep things moving smoothly so that the interlocutor can concentrate on running the interviews. Some important points to remember before the interviews begin are:

- Use only English when talking to the students.
- Give simple, clear, instructions.
- Check the student mark sheets are filled in before **all** the interviews begin
- Stand up and go over to get each pair of students.
- Remind every pair to speak up as you get them for their interview.

Interlocutors and Raters

Please use the new checklists (part 12) before **and** on the day of the interviews.



8. Room and desk set-up

Room Set up

Please set up your room during the period before the interviews. Be mindful that waiting students will be in the same room when deciding where to put the waiting student rows and the interview area. Background music can be used, but should be instrumental, and if possible be played only through the speaker(s) above the waiting students. Any music should be set at a level that does not impede the interlocutor or rater from hearing the students' answers.

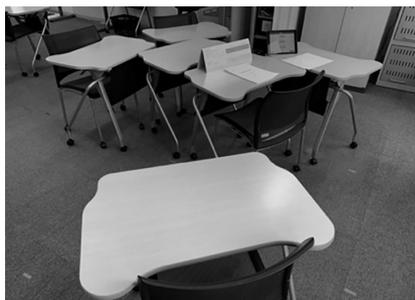
The interlocutor's desk should face the students' desks equally, and there should be a desk in-between for the Part 2 Phase 1 card. Affix A and B student signs, preferably above the students' heads on the white board behind. The interlocutor can choose to have a second desk.

The rater's desk should be as close as is possible (within two meters) but should not impede the interlocutor nor distract the students.

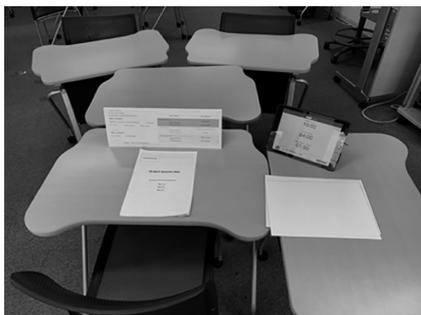
Interview Area (suggested)



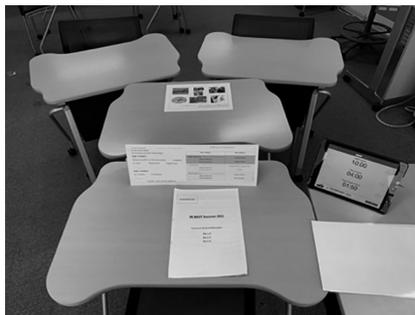
Rater / Interlocutor / Student Desks



Interlocutor Desk Set Up (suggested)



Card Placement During Part 2 Phase 1



9. Students with Special Requirements

Interlocutors and raters will be made aware of any students who have special requirements before the BEST interviews, and if necessary, will be given extra materials for these students. Students will have been consulted with by their teacher during the preceding semester to find out what their requirements are.

For students with **visual** impairments, please:

- provide them with their own Part 2 Phase 1 card on an iPad.
- allow the student a few extra seconds to zoom in and check each picture.
- confirm that the student is ready to proceed.
- if necessary, provide the student with gray-scale materials.

For students with **aural** impairments, please:

- speak at an appropriate volume.
- if applicable, seat the student in the seat with their strongest hearing side closest to the interlocutor.

For students with **physical** impairments, please be prepared to alter the seating arrangement.

10. Absent, Late, or an Odd Number of students

It is inevitable that one of the two BEST days will have an odd number of students, or that students will be delayed in coming for their interviews. There may also be students who are absent from their BEST interview.

Absent students:

- If a student is absent, the rater for that session should inform the student's teacher and the Assessment Coordinator.
- If a teacher is notified in advance of a student's absence, they should in turn contact the rater for that student's session and the Assessment Coordinator.

An odd number of students:

- **Rater:** Ask the strongest of the previous pair to stay and help. **Interlocuter:** These students are not graded a second time, and there is no need to ask them part 1 Phase 1 again.

Late students:

- **Rater:** If one member of an early pair (pairs 1~3) is absent at the beginning, keep that pair until later. Do not try to switch individuals about. If a student has still not arrived by the time the raters change rooms, please remember to inform both the interlocuter and the incoming rater.

11. Test Security

The BEST is not a 'high stakes' pass or fail assessment, but to ensure fairness and validity, we also need to keep materials secure. To ensure this, please:

- keep rooms locked until you are ready to begin the interviews.
- NEVER let students be in the rooms with materials but without a teacher.
- return the interlocuter pack to the BECC administration office at the end of your testing period(s)
- NEVER leave any assessment or grade materials in places where students could see them.

12. Interlocutor and Rater Checklists

Interlocutor

BEFORE the day:

I have...	<input checked="" type="checkbox"/>
• completed the online standardization	
• attended the in-house standardization session	
• made my name card (optional: with ordering guide cheat sheet)	
• printed my own interlocutor (holistic) rubric	
• printed my own A/B signs	
• printed out my own Interlocutor / Rater schedule	
• read the interlocutor script and familiarized myself with the running order / questions / cards	
• prepared my interview timer (timer+ is recommended)	
• found out if any of the candidates need specialized materials	
• received and signed my handbook	

On the day:

I have...	<input checked="" type="checkbox"/>
• collected my interlocutor pack	
• set up my desk(s) with the interlocutor script, timer, rubric, and name card.	
• made sure the Part 2 cards are to hand but out of sight	
• set up BGM in a way that does not interfere with my hearing the students	
• Turned my Timer+ alarms and iPad sound OFF	
• Turned Notifications off / Turned on airplane mode on my iPad	

• conducted all my interviews following the interlocutor script (including non-verbally checking with the Rater during Part 2 phase 1)	
• filled in all the student mark sheets correctly	
• returned the part 2 cards to the interlocutor pack	
• helped returned the room to 'normal'	
• returned my interlocutor pack	

• filled in my grades in the Excel doc (by 2:00pm)	
• securely stored my student mark sheets	

Rater

BEFORE the day:

I have...	<input checked="" type="checkbox"/>
• completed the online standardization	
• attended the in-house standardization session	
• made my name card	
• printed my own rater (analytic) rubric	
• printed and cut out enough student marking papers for the room I rate in first.	
• received the ordering / attendance / scoring sheet.	
• found out if any of the students need specialized materials	

On the day:

I have...	<input checked="" type="checkbox"/>
• set up my desk with my name card and rubric	
• # IF necessary: checked how to record the interviews	
• put my desk in a suitable position to be able to hear the students	
• put out the student marking sheets	
• greeted and seated the students	
• # IF necessary: put out / asked students to fill in the video consent forms	
• taken the attendance / marked any absences	
• asked students to fill in the mark sheets (model on whiteboard)	
• checked that all the mark sheets have been filled in correctly	
• # IF necessary: collected the video consent forms	
• in the case of absences, confirmed with the interlocutor what to do	
• reminded students to speak up during their interviews	
• reminded waiting students to wait quietly.	
• reminded students to leave as soon as their interviews are over	
• checked that the interlocutor is ready before calling the first pair	
• reminded each pair as they come up to speak up	
• #IF necessary: adjusted my seat position once if I cannot hear	
• non-verbally checked with the Rater during Part 2 phase 1)	
• escorted each pair back to the seating area and asked them to leave	
• left the pair ordering and grade sheets on the desk for the second rater	
• moved swiftly between rooms	
• filled in all the student scores correctly	
• helped returned the room to 'normal'	

• notified Richard and the relevant SE teacher of any absences	
• filled in my grades in the Excel doc (by 2:00pm)	
• securely stored my student grade sheets	

13. Appendices

Author's Note: All appendices have been redacted from this publication for test security purposes.

References

Cambridge Language Assessment. (2020a). *Cambridge English Qualifications A2 Key Handbook for Teachers for Exams from 2020*. Retrieved from: <https://www.cambridgeenglish.org/images/504505-a2-key-handbook-2020.pdf>

Cambridge Language Assessment. (2020b). *Cambridge English Qualifications: B1 Preliminary Handbook for Teachers for Exams from 2020*. Retrieved from: <https://www.cambridgeenglish.org/Images/168150-b1-preliminary-teachers-handbook.pdf>

Council of Europe (COE). (2018). *Common European Framework of Reference for Languages (CEFR): Learning, teaching, assessment. Companion Volume with New Descriptors*. Retrieved from: <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>

Sugg, R. & Svien, J. (2018). *Standardizing Teacher Training for CEFR-based Speaking Assessments*. Bulletin of Hiroshima Bunkyo University, Volume 53.

Further Readings

Council of Europe (COE). (2014). *The CEFR Grid for Speaking*. Retrieved from: <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806979pdf>

Council of Europe (COE). (2009). *Relating Language Examinations to the 'Common European Framework of References for Languages: Learning, Teaching, Assessment' (CEFR). A Manual*. Retrieved from: <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680667a2d>

Luoma, S. (2004) *Assessing Speaking*. Cambridge, Cambridge University Press, UK.

Talandis Jr, J. (2017) *How to Test Speaking Skills in Japan*. Alma Publishing, Japan.

All Cambridge Key and Preliminary Example Speaking Test videos are freely available at: <https://www.youtube.com/user/cambridgeenglishtv>