Questions and answers about language testing statistics:

Developing and using rubrics: Analytic or holistic?

James Dean Brown
brownj@hawaii.edu
University of Hawai'i at Mānoa

Question:

A big question in many Asian countries right now is how to make good quality rubrics for assessing oral and written English. Could you give me some tips on how to do that?

Answer:

This is the first of two columns that I will use to answer your question. In this one, I will talk about the different types of rubrics that can be used for either oral or written language output. In the next column, I will describe the steps you might take in developing a rubric, how you can decide on the categories you want to rate, and how you can approach writing the descriptors inside the cells of the rubric. In this column, I will address five central questions:

1. What is a rubric in language assessment?
2. What are analytic rubrics?
3. What are holistic rubrics?
4. What are the primary differences between analytic and holistic rubrics?
5. Where can I get more information on rubrics?

What is a rubric in language assessment?

In language testing, a *rubric* takes one of two forms as follows: (a) with language behavior categories labeled on one dimension of a rectangular matrix and scores labeled on the other dimension; in the cells of the matrix each score level is described for each category in terms of the expected language performances for that score level and category (e.g., see Table 1) and (b) with scores labeled along one dimension of the matrix and descriptors supplied next to each score that describe the language behaviors expected at each score level in terms of expected language performances (see Table 2).

Among other things, rubrics can be used for scoring students' language abilities or for giving students feedback on their language learning progress or achievement in learning those language abilities. In particular, rubrics provide useful tools for assessing students' abilities to use their productive language skills of speaking or writing, or to use their productive and receptive skills (listening and reading) in interactions with each other. In curricula that include portfolios, or task-based activities (like writing an email, filling out a form, surviving a job interview, doing a presentation, etc.), rubrics can prove especially useful.

What are analytic rubrics?

The rubric shown in Table 1 is an example of an *analytic rubric*, in this case one developed for giving feedback and scoring written letters. Notice that the language categories are labeled on the left for each of the rows and that the possible scores are labeled across the top for each column. The categories in this

case for writing a letter are: Salutation & Closing; Length; Grammar & Spelling; Capitalization & Punctuation; and Neatness. The scores in this case are 4, 3, 2, and 1. The categories that I decided to use here seemed important to me for teaching good letter writing, but a different teacher might think that totally different categories should be used. The decision of what categories to use in an analytic rubric should be based on categories that: the teacher thinks are important; the teacher wants the students to focus and work on; the teacher wants to give students feedback on; and perhaps, the teacher wants to score and/or grade. Thus, the categories should be well-thought-out and should remain open to revision if the teacher changes her mind about what is important.

Table 1

Analytic Rubic for Scoring Written Letters[1]

Category	4	3	2	1
Salutation & Closing	Salutation and closing have no errors in capitalization and punctuation.	Salutation and closing have 1-2 errors in capitalization and punctuation.	Salutation and closing have 3 or more errors in capitalization and punctuation.	Salutation and/or closing are missing.
Length	The letter is 10 or more sentences.	The letter is 8-9 sentences.	The letter is 5-7 sentences.	The letter is less than 5 sentences.
Grammar & Spelling (conventions)	Writer makes no errors in grammar or spelling.	Writer makes 1-2 errors in grammar and/or spelling.	Writer makes 3-4 errors in grammar and/or spelling	Writer makes more than 4 errors in grammar and/or spelling.
Capitalization & Punctuation	Writer makes no errors in capitalization and punctuation.	Writer makes 2-4 errors in capitalization and punctuation.	Writer makes 4-6 errors in capitalization and punctuation.	Writer makes more than 6 errors in capitalization and punctuation.
Neatness	Letter is typed, clean, not wrinkled, and is easy to read with no distracting error corrections. It was done with pride.	Letter is neatly hand-written, clean, not wrinkled, and is easy to read with no distracting error corrections. It was done with care.	Letter is typed and is crumpled or slightly stained. It may have 1-2 distracting error corrections. It was done with some care.	Letter is typed and looks like it had been shoved in a pocket or locker. It may have several distracting error corrections. It looks like it was done in a hurry or stored improperly.

Notice then that each score level for each category is described as clearly as possible. For example, to get a 4 for Salutation & Closing, the student would have to do so with "no errors in capitalization and punctuation." Such descriptions are very personal. This being what I think a student should do to get a 4. Again, another teacher might describe this quite differently. And that of course is fine.

The rubric shown in Table 2 is also an example of an analytic rubric but for giving feedback on and scoring student oral presentations. This time, the language categories are labeled across the top for each column and the possible scores are labeled down the left side for each of the rows. The categories in this case for student presentations are: Preparedness; Content; Enthusiasm; Speaks Clearly; and Posture &

---

[1] Note that the rubrics in Tables 1 & 2 were generated online using Rubistar (http://rubistar.4teachers.org/); the rubrics in Tables 3 & 4 were adapted from Tables 1 & 2, respectively, by using *Excel[tm]* to consolidate the text.

Eye Contact. The scores in this case are also 4, 3, 2, and 1. The categories that I decided to use here seemed important to me for preparing students to do good presentations in class, but again, a different teacher might think that completely different categories should be used, and that is as it should be.

Table 2

Analytic Rubic for Scoring Student Oral Presentations

	Preparedness	**Content**	**Enthusiasm**	**Speaks Clearly**	**Posture & Eye Contact**
4	Student is completely prepared and has obviously rehearsed.	Shows a full understanding of the topic.	Facial expressions and body language generate a strong interest and enthusiasm about the topic in others.	Speaks clearly and distinctly all (100-95%) the time, and mispronounces no words.	Stands up straight, looks relaxed and confident. Establishes eye contact with everyone in the room during the presentation.
3	Student seems pretty prepared but might have needed a couple more rehearsals.	Shows a good understanding of the topic.	Facial expressions and body language sometimes generate a strong interest and enthusiasm about the topic in others.	Speaks clearly and distinctly all (100-95%) the time, but mispronounces one word.	Stands up straight and establishes eye contact with everyone in the room during the presentation.
2	The student is somewhat prepared, but it is clear that rehearsal was lacking.	Shows a good understanding of parts of the topic.	Facial expressions and body language are used to try to generate enthusiasm, but seem somewhat faked.	Speaks clearly and distinctly most (94-85%) of the time. Mispronounces no more than one word.	Sometimes stands up straight and establishes eye contact.
1	Student does not seem at all prepared to present.	Does not seem to understand the topic very well.	Very little use of facial expressions or body language. Did not generate much interest in topic being presented.	Often mumbles or cannot be understood OR mispronounces more than one word.	Slouches and/or does not look at people during the presentation.

What are holistic rubrics?

The rubric shown in Table 3 is an example of a *holistic rubric* for scoring written letters. Notice that, this time, there are no language categories though the possible scores are labeled down the left side for each of the rows. The categories in this case are subsumed in the descriptions, where you will note that salutation & closing, length, grammar & spelling, capitalization & punctuation, and neatness are all touched on for each score level. Indeed, a quick comparison will show the reader that the words at each level were simply block copied from those in the first column of Table 1.

Similarly, the rubric shown in Table 4 is an example of a holistic rubric for student oral presentations. Notice again there are no language categories though the possible scores are labeled down the left side for each of the rows. Again, the categories of preparedness, content, enthusiasm, speaks clearly, and posture & eye contact are all touched on for each score level in the descriptions. And again, a quick comparison

will show the reader that the words at each level were simply block copied from those in the first row of Table 2.

Table 3
Holistic Version of the Rubic for Scoring Written Letters

Score	Description
4	Salutation and closing have no errors in capitalization and punctuation. The letter is 10 or more sentences. Writer makes no errors in grammar or spelling. Writer makes no errors in capitalization and punctuation. Letter is typed, clean, not wrinkled, and is easy to read with no distracting error corrections. It was done with pride.
3	Salutation and closing have 1-2 errors in capitalization and punctuation. The letter is 8-9 sentences. Writer makes 1-2 errors in grammar and/or spelling. Writer makes 2-4 errors in capitalization and punctuation. Letter is neatly hand-written, clean, not wrinkled, and is easy to read with no distracting error corrections. It was done with care.
2	Salutation and closing have 3 or more errors in capitalization and punctuation. The letter is 5-7 sentences. Writer makes 3-4 errors in grammar and/or spelling. Writer makes 4-6 errors in capitalization and punctuation. Letter is typed and is crumpled or slightly stained. It may have 1-2 distracting error corrections. It was done with some care.
1	Salutation and/or closing are missing. The letter is less than 5 sentences. Writer makes more than 4 errors in grammar and/or spelling. Writer makes more than 6 errors in capitalization and punctuation. Letter is typed and looks like it had been shoved in a pocket or locker. It may have several distracting error corrections. It looks like it was done in a hurry or stored improperly.

Table 4
Holistic Version of the Rubic for Scoring Student Oral Presentations

Score	Description
4	Student is completely prepared and has obviously rehearsed. Shows a full understanding of the topic. Facial expressions and body language generate a strong interest and enthusiasm about the topic in others. Speaks clearly and distinctly all (100-95%) the time, and mispronounces no words. Stands up straight, looks relaxed and confident. Establishes eye contact with everyone in the room during the presentation.
3	Student seems pretty prepared but might have needed a couple more rehearsals. Shows a good understanding of the topic. Facial expressions and body language sometimes generate a strong interest and enthusiasm about the topic in others. Speaks clearly and distinctly all (100-95%) the time, but mispronounces one word. Stands up straight and establishes eye contact with everyone in the room during the presentation.
2	The student is somewhat prepared, but it is clear that rehearsal was lacking. Shows a good understanding of parts of the topic. Facial expressions and body language are used to try to generate enthusiasm, but seem somewhat faked. Speaks clearly and distinctly most (94-85%) of the time. Mispronounces no more than one word. Sometimes stands up straight and establishes eye contact.
1	Student does not seem at all prepared to present. Does not seem to understand the topic very well. Very little use of facial expressions or body language. Did not generate much interest in topic being presented. Often mumbles or cannot be understood OR mispronounces more than one word. Slouches and/or does not look at people during the presentation.

What are the primary differences between analytic and holistic rubrics?

Initially, you may have thought that the differences between holistic and analytic rubrics were relatively superficial: more about differences in format than about the content of the grids. Indeed, the content of

Tables 1 and 3 is exactly the same. I simply block-copied the words (describing Salutation & Closing; Length; Grammar & Spelling; Capitalization & Punctuation; and Neatness) in the row for a score of 4 from Table 1 and pasted them into the to the descriptor for a score of 4 in Table 3. The same is true for each of the subsequent rows in Table 1. The result is certainly a difference in format, but that difference fundamentally changes both the usefulness and purpose of the rubric. In fact, choosing to develop a holistic or analytic rubric will affect everything else in the scoring procedures. It is therefore probably a good idea to decide on the format early and for sound reasons. What might those reasons be?

Consider that *holistic scoring* provides a single global scale that will produce a single general rating for each examinee's productive language sample. Thus, holistic scoring will, by definition, produce a single score, which makes it quicker and easier to apply than analytic scoring. However, holistic scoring is not very useful for providing itemized feedback to examinees. Hence, holistic scoring is most often used for overall proficiency testing to support institutional or programmatic admissions decisions or for placement testing to determine what levels of language different examinees should study. In both cases, time is usually short and itemized feedback to examinees is not necessary.

In contrast, *analytic scoring* provides multiple scores for different aspects of each examinee's productive language sample. Thus, analytic scoring is, by definition, better for giving feedback on multiple aspects of the examinees' language performances, aspects that the teacher deems important to emphasize. Analytic scoring is therefore more difficult and time consuming to do, but teachers often feel that the effort is worthwhile because it allows them to provide very useful itemized or detailed feedback to examinees. Hence, analytic scoring is most often used in classroom assessment for diagnostic, progress, and/or achievement testing.

Where can I get more information on rubrics?

General education articles and books about the development and use of rubrics include Arter and McTighe (2001), Campbell Hill and Ekey (2010), Glickman-Bond and Rose (2006), and Hutson-Nechkash (2003), Popham (1997), Mertler (2001), Moskal (2000), and Tierney and Simon (2004).

Language testing books that included at least one chapter that covers rubrics include Brown (2005), Brown and Hudson (2002), and Buttner (2007). Many articles directed at language testers center on technical aspects of rubric design, development, and analysis, especially rubrics used in large-scale, high-stakes testing. However, Upshur and Turner (1995) describe strategies for designing rubrics in language assessment, and a number of authors describe the development and use of rubrics in language classrooms for: *language for specific purposes* (for example, see Arnold, 1998; Blankmann, 1998; Ho, 1998; Johnson, 1998; Russ, 1998; and Shimazaki, 1998); *speaking* (see Luoma, 2004, pp. 59-95; pp. 226-255); and *writing* and *portfolios* (see Weigle, 2002, pp. 140-171, 190-196, 222-227). In my humble opinion, the best single book on developing, administering, and analyzing rubrics for language assessment purposes is the collection of articles provided in Brown (2012).

Conclusion

In direct answer to your question, "A big question in Asian countries right now is how to make good quality rubrics for assessing oral and written English. Could you give me some tips on how to do that?" Here I have tried to give you "tips" in the form a first step which involves thinking through which option you want to use: a holistic rurbric or an analytic rubric. In the next column, I will explain the steps you might take in developing a rubric, how you can decide on the categories you want to rate, and how you can approach writing the descriptors inside the cells of the rubric.

I hope this column and the next one together will address your question and provide you with the information you will need to at least get started in developing and using rubrics for assessing oral and written English and for giving your students useful feedback.

References

Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Thousand Oaks, CA: Corwin/Sage.

Blankmann, J. (2013). Oral presentations: How did I do? In J. D. Brown (Ed.), *New ways of classroom assessment, revised* (pp. 117-119). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment* (New edition). New York: McGraw-Hill.

Brown, J. D. (Ed.) (2012). *Developing, using, and analyzing rubrics in language assessment with case studies in Asian and Pacific languages.* Honolulu, HI: National Foreign Languages Resource Center.

Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing.* Cambridge: Cambridge University.

Buttner, A. (2007). *Activities, games, assessment strategies, and rubrics: For the foreign language classroom.* Larchmont, NY: Eye on Education.

Campbell Hill, B., & Ekey, C. (2010). The next-step guide to enriching classroom environments: Rubrics and resources for self-evaluation and goal setting for literacy coaches, principals, and teacher study groups, K-6. Portsmouth, NH: Heinemann.

Glickman-Bond, J., & Rose, K. (2006). *Creating and using rubrics in today's classrooms: A practical guide.* Norwood, MA: Christopher-Gordon.

Ho, B. (1998). How well did I communicate? In J. D. Brown (Ed.), *New ways of classroom assessment* (pp. 97-98). Washington, DC: Teachers of English to Speakers of Other Languages.

Hutson-Nechkash, P. (2003). *Help me write: Frames and rubrics for classroom writing success.* Greenville, SC: Thinking Publications.

Johnson, J. (2013). So how did you like my presentation. In J. D. Brown (Ed.), *New ways of classroom assessment, revised* (pp. 95-97). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Luoma, S. (2004). *Assessing speaking*. Cambridge, England: Cambridge University.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation,7*(25). Retrieved October 26, 2010 from http://PAREonline.net/getvn.asp?v=7&n=25 .

Moskal, B. M. (2000). Scoring rubrics: what, when and how? *Practical Assessment, Research & Evaluation, 7*(3). Retrieved October 26, 2010 from http://PAREonline.net/getvn.asp?v=7&n=3 .

Paltridge, B. (2013). Observation, feedback, and individual goal setting. In J. D. Brown (Ed.), *New ways of classroom assessment, revised* (pp. 185-189). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Popham, W. J. (1997). What's wrong – and what's right – with rubrics. *Educational Leadership, 55*(2), 72-75.

Russ, R. (2013). Let me explain. In J. D. Brown (Ed.), *New ways of classroom assessment, revised* (pp. 160-162). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Shimazaki, R. (2013). Guess what my favorite animal is. In J. D. Brown (Ed.), *New ways of classroom assessment, revised* (pp. 356-357). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*, 9(2). Retrieved January 9, 2011 from http://PAREonline.net/getvn.asp?v=9&n=2.

Upshur, J. A., & Turner, C. E. (1995). Constructing rating scales for second language tests. *ELT Journal, 49*(1), 3-12.

Weigle, S. C. (2002). *Assessing writing.* Cambridge, England: Cambridge University.

Where to submit questions:

Your question can remain anonymous if you so desire. Please submit questions for this column to the following e-mail or snail-mail addresses:

brownj@hawaii.edu.

JD Brown
Department of Second Language Studies University of Hawai'i at Mānoa
1890 East-West Road
Honolulu, HI 96822 USA