

Evaluating the Usefulness and Properties of a Subjective Assessment of Brazilian Portuguese¹

Antônio R.M. Simões and Elena C. Papanastasiou
University of Kansas

Abstract: This is an evaluation of a Brazilian self-assessment test. All questions on the test guide the students to evaluate themselves on their linguistic and socio-cultural skills in Brazilian Portuguese. The main points discussed in this article are (1) an item analysis, (2) the reliability of the test to determine the consistency of the results obtained by the instrument, (3) the construct validity, and (4) the creation of cutscores. Our analysis of the BP self-assessment was supported by the use of descriptive statistics, by a factorial analysis and by a reliability test to determine the psychometrics characteristics of the test. The major claim in this study is that this test is useful, valid and reliable, if used appropriately, especially with an audience of motivated students such as students going abroad, instead of students who take language classes only as a requirement in their school program.

Key Words: Brazilian Portuguese, cutscores, foreign language, item analysis, placement, psychometric properties, reliability, self-assessment, testing, validity

1. Introduction

Although self-assessment has received more attention in recent years (Oskarsson 1997), it is an area of testing that has not been adequately researched yet because of the subjective nature of such tests. In non-competitive situations, self-assessments could be valuable tools in the area of language testing if their properties and usefulness are adequately determined. Therefore, the purpose of this paper is to determine the qualities, usefulness, and psychometric properties of a self-assessment tool for the Brazilian Portuguese (BP) language. This paper will also provide a context and an example in which such tests can be used.

Among the studies that exist on self-assessment, we find LeBlanc and Painchau's study (1985) where they explored the use of a self-assessment questionnaire as a placement instrument. In their study, two hundred students studying either French or English as a second language were randomly selected at the University of Ottawa to participate in the study. What they did was to ask the students to fill out a self-assessment questionnaire prior to taking a proficiency test. Their data showed a high (.53) correlation between the two tests. As a result of their study, the administration decided to use the self-assessment questionnaire as their main placement instrument. However, the authors warn that self-assessment tests should only be used when the students have nothing to lose or gain from the results. Other authors, e.g. Oskarsson (1978, 1997), also favor the use of self-assessment. In his 1978 book, Oskarsson argues that students can accurately self-evaluate their own language capabilities, especially when asked to judge specific tasks rather than overall language competence.

Pierce, Swain and Hart (1993) have studied the self-assessment of French proficiency of approximately 500 Grade 8 students in "early" and "middle" programs in Toronto, Canada. Their study analyzed the self-assessment of junior-high students in comparison with two self-assessment benchmarks, the perceived language proficiency of Francophone peers and the difficulty represented by everyday tasks in French. One of their basic questions was: Are the students evaluating themselves in comparison to the benchmark of Francophone peers or with the

benchmark of their difficulty in completing specific tasks? Their study was designed to compare the results of both benchmarks. These authors also advise caution in the use of self-assessments because one of the main results of their investigation is that only when self-assessment tools have specific and focused tasks the correlation between self-assessments and objective assessments will be high.

Therefore, the studies described above are very useful guides for the improvement of self-assessment techniques. They confirm the usefulness of self-evaluation as a testing tool, provided that test creators are cautious about the preparation of test items. The BP self-assessment has been prepared with caution. Through this study, we will show that as long as a subjective test is statistically reliable, valid and has well constructed cutscores for placing students in appropriate language levels, it can become an invaluable tool if used in the appropriate context.

For the purpose of this study, it should be noted that this subjective test for BP is currently intended to be used as an independent non-competitive test or as a supplement for making placement decisions for students who will be attending a foreign language program in BP. It is important to note that we are dealing with a population of motivated students and not the more common type of college students who take language classes because it is a requirement.

The self-assessment for BP used in this study was administered for the first time in 2001 as a placement tool of students who participated in the University of Kansas Summer Language and Culture Institute, in Brazil. Before the creation of the self-assessment, the Brazil program used to give an objective test to its students, which took a lot of time and energy to be graded and to place the students in the appropriate language levels. However, even after the student's placement, quite a number of them would regularly ask to be changed to a higher or lower language level group. With the use of the self-assessment, this trend started to disappear. With the first use of the current self-assessment, all students stayed properly placed, except for two out of a total of thirty students. Those two students wanted to move from the advanced to the intermediate level. However, these two students were clearly advanced as we attested from their later performance in the program. In our opinion, these students simply wanted to be in a less demanding classroom program. Consequently, we can still argue that the self-assessment performed its function to place them in their appropriate group.

Therefore, the purpose of this study was to determine if the BP subjective assessment actually has good psychometric properties so that it can be useful WHEN used properly in a non-competitive situation and it is quite clear to us that it can.

2.0 Methodology

The purpose of this section is to describe the self-assessment tool, which is the focus of this study, the way it was administered, as well as a description of the statistical analyses that were performed on this test.

2.1 The self-assessment and its application

The 48 questions that were included in the self-assessment dealt with the content areas of overall speaking skills, precision in speaking, culture and civilization, listening, reading, and writing. The questions on the test also guided the students to evaluate themselves on their socio-cultural skills in addition to their linguistics skills in BP. More specifically, the students had to subjectively assess themselves on if they believed that they knew the information on each of the 48 questions. A student who believed that they knew the information contained in all 48 questions would obtain a total score of 48. A student who believed that they knew none of the information on the test would obtain a score of 0 out of 48. All questions required the test users to answer "yes" or "no." The choice of a neutral "not sure" or "I don't know" option was purposely left out of the questionnaire since we wanted to eliminate the pattern of behaviors where students conveniently choose the neutral option without putting much thought into it.

All items in the BP self-assessment have been written by one person, and reviewed by several colleagues and graduate students in Brazil and in the US, as well as by two language specialists in the US. The complete self-assessment and its rationale can be viewed at the URL <www.ukans.edu/~brasilis> at the link "Brazilian Portuguese testing materials." Appendix 1 contains all 48 question items included in the questionnaire, but not the complete test, which also includes background additional questions such as the student's age group and address. The subjective test was administered to 124 college/university students and teachers in the US and also in Brazil, between February 2001 and October 2001. All participants took the test online from different campuses in the US and Brazil.

2.2 Statistical analyses

The first stage of the investigation included the use of descriptive statistics to describe the performance of the students on the test, while the second stage included an examination of the psychometric properties of the test.

As a first step, this investigation examined the results obtained from an item analysis. This was done to determine the quality of the items used in the instrument, as well as how well each item performed in discriminating the students based on their subjective assessment of their knowledge and skills (Crocker and Algina 1986, Mehrens and Lehman 1991). Following, a reliability test was performed to determine the consistency of the results obtained by the instrument (McMillan and Wergin 2002).

The construct validity of the subjective testing instrument was also examined through a factor analysis to determine the extent to which it measured what it was supposed to measure (Pedhazur and Pedhazur Schmelkin 1991). However, many of the popular statistical packages such as SPSS and SAS can only factor analyze data that are continuous in nature. This was a problem because the data obtained from this study were dichotomous, since the students could respond to the self-assessment by choosing the answers of either yes or no. Trying to perform a factor analysis with packages such as SPSS or SAS would only lead to misleading conclusions since a major assumption about the format of the data would clearly be violated. Thus, to perform the factor analysis of the data, a specialized factor analysis software package for dichotomous data called Microfact (Waller 1995) was used. This software can properly factor analyze dichotomous data since it computes tetrachoric correlations for such data, rather than Pearson product moment correlations that are used for continuous data (Waller 1995). In addition, a varimax rotation was selected as a factor rotation method, to help with the interpretation of each factor.

Finally, an attempt was also made to describe the creation and use of cutscores that could be used for the placement of students in the proper foreign language program.

3.0 Results and discussion

A summary and discussion of the study's results will be presented next. These results confirm the usefulness, validity and reliability of the current test, as well as some additional psychometric properties. Therefore, the main points discussed below will be (1) an item analysis, (2) the reliability of the test to determine the consistency of the results obtained by the instrument, (3) the construct validity of the instrument, and (4) the creation of cutscores.

3.1 Item analysis

The characteristics of the test items are presented in Table 1, which describes the difficulty level of each of the 48 questions on the test. The difficulty level is determined by the proportion of people that responded that they knew the information of each question. A high proportion indicates that the question is easy since many students have responded that they knew that information, while a low proportion indicates that the question was difficult for the students.

For this test, question 3 (superior) that stated “Do you feel you have the appropriate language skills to introduce a celebrity, in Portuguese, at a formal ceremony? This would require a native knowledge of social and cultural protocols among Brazilians as well as no grammar mistakes,” was the most difficult one since only 14 students (11.3%) responded that they knew the information of that question. In contrast, question 21 (basic) that stated “Could you greet a Brazilian friend with simple expressions equivalent to the English expression ‘How are you?’ in a casual setting?” was the easiest one since 122 out of the 124 students (98.4%) responded that they knew its answer.

A point-biserial correlation coefficient was also calculated for each item on this test. The purpose of this correlation is to examine how closely the performance on each test item resembles the student’s total test score. The results of this analysis that are presented in Table 2 were very positive since all 48 items were significantly and positively correlated to the total test score at the 0.05 level. Since this is true, there were no questions to which lower ability students would respond that they knew the answer to a question, while higher ability students would respond that they did not know the answer to that question. This indicates that all questions on the test functioned properly as desired.

3.2 Test reliability

The reliability of the test scores was examined to determine the consistency of the results obtained by the instrument. This is done by measuring the amount of error that exists in the scores obtained by the instrument. These random errors can occur either because of the individuals taking the tests, or because of the instrument itself. For example, if a student were administered the same instrument twice, it is likely that they would obtain different scores from each administration. This could occur because of reasons such as fatigue, motivation and attitudes at the moment in which the instrument is administered. Errors because of the instrument could occur if there were items that were not worded properly or clearly. So the reliability estimate of the test scores is useful for determining how reliable and how consistent the results obtained by the instrument can be. The Cronbach’s alpha estimate was used to obtain the reliability of this test, and the reliability that was obtained was 0.92. This indicated that the standard error of the score estimates is low, and that the subjective test is highly reliable since the value of the reliability is very close to 1.0.

3.3 The construct validity

The validity of the subjective testing instrument was also examined to determine the extent to which it measured what it was supposed to measure. The type of validity evidence that was used to obtain this information was that of construct validity. This was done to examine what are the underlying constructs that are measured by the instrument.

Therefore, a factor analysis was performed to determine the amount of underlying factors that could account for the variation of the 48 items on the test. In addition, a varimax rotation was selected as a factor rotation method, to help with the interpretation of each factor, as seen in Tables 3 and 4.

The factor analysis that was performed showed that the assessment contained five significant factors in the self-assessment, which were labeled as factor-1—comprehension (listening and reading); factor-2—culture and civilization; factor-3—production (writing and speaking); factor-4—miscellaneous; factor-5—pronunciation specifics. These five factors altogether represent approximately 27% of the variance of the 48 items, a result that attests to the significance of the five factors and the validity of the construct of the test.

An analysis of the content of items based on Appendix 1 and Table 3 has revealed that some of the items repeat certain words and string of words that might bias test users. The information in parentheses following each question in Appendix 1 gives an idea of the original design of the

self-assessment, which was intended to have six factors as follows, (1) culture and civilization, (2) overall oral production (3) oral comprehension, (4) written production, (5) written comprehension, and (6) specific oral production.

Given the results in Table 3 and 4, we asked ourselves why the factorial analysis revealed a number of factors and groups different from what they were intended to be. What might have influenced these results is the use of repeated words or sequence of words found in different questions under different factors. By reading the contents of the questions, we can observe that a word and its derivatives, e.g. "explain/ing," appear in the questions of different language skills, i.e. speaking, listening, writing and reading. The use of a dictionary is also mentioned repeatedly through different factors and levels. This repetition of words and instructions might have some influence on the pattern which the test takers responded to the different questions on the test.

The results from this session attest to the validity evidence of the BP self-assessment instrument. However, more work can always be done to increase its validity. This can be attained by including additional questions in each factor, and by pilot testing these questions on more students.

3.4 Creation of cutscores

To create cutscores we can use a distribution of test results of the 124 students shown in Figure 1, combined with Table 5, to place the students in different groups. From a distribution of this kind we can place students in many different groups, depending on the needs of a language program. In other words, by studying the data distribution, we could for example divide the 124 students in groups of five, six, seven or more. For instance, given the distribution in Figure 1, it would be possible to create eight groups of students with cutscores at 1–16 (group 1), 17–24 (group 2), 25–27 (group 3), 28–31 (group 4), 32–36 (group 5), 37–40 (group 6), 41–45 (group 7) and 46–48 (group 8).

During the University of Kansas summer language program in 2001, cutscores were used to determine the placement of the 30 students who participated in that program. KU's program needs to divide these students in three groups. Therefore, creation of cutscores using the subjective test during the summer of 2001, allowed us to create three groups to fit the goals of the summer program. The placement was based on the following cutscores: group 1, basic, 0–29; group 2, intermediate, 30–41; and group 3, advanced, 42–48. It is a fact that the creation of cutscores is always an arbitrary and subjective process. However, although the cutscores were created subjectively, they still provided help to the program to realize its goals.

Before 2001, preliminary versions of the current self-assessment were given twice as a pilot test for the current version. After the students abroad used the current self-assessment online, we realized that the students were in fact very good at placing themselves in a given level. It is important to note that the students in our program come from several US campuses while only half-dozen out of thirty summer school students came from the University of Kansas.

As shown in Figure 1, the score distribution was bimodal and slightly negatively skewed. The two modes of the distribution were 29 and 33 with a frequency of 10 students on each mode/score. As shown in Table 5, the average score obtained by all 124 students was 31.2, with a standard deviation of 8.3, while the median score of the distribution was 32. The range of scores was 46, since the scores ranged from a 2 to a 48.

4. Limitations and directions for further research

The results of this study clearly depict the usefulness of this self-assessment test for BP. The data that we have collected have shown that this self-assessment is not only a useful language tool with valid and reliable psychometric properties, but it has also paved the ground for other testing material and related research.

Although this study has revealed many positive results, additional research can be done to

increase our understanding of self-assessments. One of the next steps of this study will be to cross-validate its results for a stronger generalization of our claims. This cross-validation would require the test to be administered to an even larger group of students in the future to determine if similar psychometric properties will be found with a different sample of students.

Another limitation of this study is that we were not able to study the predictive validity of this self-assessment. Examining this type of validity is essential to this assessment since it will allow us to determine how this assessment can predict the placement of the students into BP language levels, as well as how well it will predict the student's overall performance on the objective test and in these BP courses. However, the small sample size of the students that attended this summer school program, as well as the current and temporary lack of an objective test did not allow us to examine this type of validity evidence. In addition, although there are a few objective tests for Brazilian Portuguese, they do not cover the same areas of interest for adequate comparison. Consequently, after the creation of an objective test for Brazilian Portuguese, this type of validity will be examined.

Another direction for further research that has evolved with respect to the use of the self-assessments for non-competitive situations has to do with the area of adaptive testing. Given that one of the purposes of this subjective test is to place students in one of seven levels of BP, this classification will then be used as an initial ability estimate for an objective adaptive test that will be administered to the students after they have completed the subjective test. The assignment of initial ability estimates can be very beneficial for small-scale adaptive tests. First, the items of average difficulty will not be overexposed to the students by continuously administering them at the beginning of the test. Second, the estimation of the ability estimates will be more accurate since there will be more items that will be targeted directly to an examinee's ability. Finally, this would lead to the third advantage, which is that the standard error of the ability estimates will be smaller, while the information obtained by the test would be larger.

5. Conclusion

The main purpose of this investigation was to find out whether or not a BP self-assessment can be a useful tool in BP language programs and to study its properties. The creation of a language tool of this nature requires more than one phase to become adequately operative. Before going into a second phase of the cross validation of these results however, additional data needs to be obtained from a larger sample of students.

The preliminary results from this study have made us confident but also cautious about how to proceed. A similar alert has been seen in former studies about self-assessment (Oscarson 1978; Pierce, Swain and Hart 1993), which state that self-assessments are likely to be more effective when they have specific tasks. We feel confident about the current Brazilian Portuguese self-assessment not only because of its positive results, but also because the content of its items are generally characterized by the specificity of its questions, namely its tasks.

Although there might be other uses for this test, in its current stage it is intended to be used as an independent non-competitive test or as a supplement for making placement decisions for students who will be attending a language program in BP. Our goal becomes easier to reach and our claims stronger if one takes into account that we have aimed this test for an audience of motivated students such as students going abroad, instead of students who take language classes only as a requirement in their school program.

This investigation has provided us with the necessary ground for other testing research and also a number of invaluable information useful in the design of foreign language programs. In our opinion, this is so, because the more experience we accumulate with the use of adequate self-assessments, the more we will be able to develop and implement clearer goals to be achieved in language curricula.

NOTE

¹This research was made possible thanks to a U.S. Department of Education Title VI grant, through the program International Research and Studies.

WORKS CITED

- American Council on the Teaching of Foreign Languages* (ACTFL), Edited by Kathryn Buck, and contributing editors Heidi Byrnes and Irene Thompson. Yonkers, NY: ACTFL Inc. 1989.
- Bailey, Kathleen M. *Learning about language assessment: Dilemmas, decisions, and directions*. Pacific Grove, CA: Heinle and Heinle Publishers, 1998.
- Blanche, Patrick and Barbara J. Merino. "Self-assessment of foreign-language skills: Implications for teachers and researchers." *Language Learning: A Journal of Applied Linguistics* 39-3 (1989): 313-340.
- Crocker, Linda M. and James Algina. *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers, 1986.
- LeBlanc, Raymond and Gisèle Painchaud. Self-assessment as a second language placement instrument. *TESOL Quarterly* 19-4 (December, 1985): 673-687.
- McMillan, James H. and Jon F. Wergin. *Understanding and evaluating educational research*. Upper Saddle River, NJ: Merrill Prentice Hall, 2002.
- Mehrens, William A. and Irvin J. Lehman. *Measurement and evaluation in education and psychology*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers, 1991.
- Oscarsson, Mats. *Approaches to self-assessment in foreign language learning*. Oxford: Pergamon Press, 1978.
- . "Self-Assessment of Foreign and Second Language Proficiency." *The Encyclopedia of Language and Education*. Ed. Caroline Clapham and David Corson. Dordrecht, Netherlands: Kluwer Academic Publishers 7 (1997): 175-187.
- Pedhazur, Elazar J. and Liora Pedhazur Scmelkin. *Measurement, design and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 1991.
- Pierce, Bonny N., Merrill Swain, and Doug Hart. "Self-assessment, French immersion, and the locus of control." *Applied Linguistics*, 14-1 (March, 1993): 25-42.
- Simões, Antônio R.M. *Brazilian Portuguese Self-assessment Test*. Internet URL: <www.ukans.edu/~brasilis>, 2000.
- Waller, Niels G. *MicroFACT 1.1 — A microcomputer factor analysis program for ordered polytomous data and mainframe size problems*. St. Paul, MN: Assessment Systems Corp., 1995.

Appendix 1

Self-Assessment Exam for Brazilian Portuguese

© 2000 Antônio R.M. Simões

These are the 48 items of the self-assessment in random order. The labels and numbers in parentheses after each question correspond to the proposed language areas and levels of the questions. Each level has two sublevels, namely Basic 1, Basic 2, Intermediate 1, Intermediate 2, Advanced 1, Advanced 2, Superior 1 and Superior 2. For example, (Cult/Civ-Adv 1) means "area of culture and civilization, at the advanced, first level."

The test users are required to answer "yes" or "no" to each question item. The complete test on the internet has additional information about the participant's personal information, the test rationale, instructions regarding test taking, and other additional information.

1. If you were the boss in an enterprise and just employed a secretary who only spoke Portuguese, would you be able to explain to the secretary the details of her functions and her terms of employment with your company? (OverallSpeaking-Superior 2)
2. Did you know that Brazil has had three capitals in the history of the country: Salvador, Rio de Janeiro and Brasília? (Cult/Civ-Adv 2)
3. Do you feel you have the appropriate language skills to introduce a celebrity, in Portuguese, at a formal ceremony? This would require a native knowledge of social and cultural protocols among Brazilians as well as no grammar mistakes. (OverallSpeaking-Superior 1)
4. In your hotel, the receptionist only speaks Brazilian Portuguese. The receptionist is explaining to a friend of yours who only speaks English, the rules of the hotel and other helpful information about the city. Do you think that you can understand the explanation and translate it to your friend? (Listening-Adv 1)
5. Can you understand most of the text below (90% or more)?
 "Com relação ao novo sucesso de bilheteria, Gladiador, neste excelente filme os gladiadores morrem como moscas, enfrentam animais selvagens e soldados montados em carruagens. Porém, na realidade, gladiadores profissionais eram muito valorizados para morrer assim. Em geral, a luta era interrompida quando um deles desistia ou se feria e as feras eram usadas apenas para devorar os condenados. Também nesse filme, o imperador Marco Aurélio era contra os

espetáculos de violência e suspendeu as lutas. Porém o fato histórico é que Marco Aurélio, como todos os outros imperadores incentivavam o festival de pancadaria no Coliseu” (Reading-Adv 2)

6. Are you familiar with the Portuguese expression “Dar um jeitinho”? (Cult/Civ-Interm 1)
7. “Vocês não trouxeram ovos hoje, então tomem um caqui para atirar.” (Mário Covas). Do you understand most of this sentence (90% or more)? (Reading-Adv 1)
8. Have you read or studied from any source that the racial mixture that resulted in the true make up of Brazilian people consists of the Indians, the African and the Portuguese, especially from the center toward the north of Brazil? (Cult/Civ-Superior 2)
9. If you pronounce the “-m”s in word final position in Brazilian Portuguese like you would say the -m in the English word “atom,” that is considered incorrect. Do you think that your pronunciation of “-m”s in words like *também, um*, etc. is correct? (PronunciationSpecifics-Interm 1)
10. In your pronunciation of Brazilian Portuguese can you distinguish *v* and *b* in words like *votar* and *botar*? (PronunciationSpecifics-Basic 1)
11. You were chosen as the keynote speaker at a club luncheon and you have to speak to an audience of non-English speaking Brazilians. Would you be able to write an error-free, well-written speech for this luncheon on your own? (You may use reference grammars, dictionaries, etc.) (Writing-Superior 2)
12. Did you know that Brazil was discovered in 1500? (Cult/Civ-Basic 2)
13. The issue of globalization and its effects on developing countries is being discussed in Brazilian Portuguese in a conference. Do you think you would understand completely the discussion? (Listening-Superior 1)
14. Suppose someone left you a brief, simple note with the following list of items to buy: arroz, feijão, farinha, trigo, açúcar, café, pão-doce, pão-de-sal, leite, manteiga, couve, maçã, cenoura, alface, manga-espada, abacaxi, uva, coco verde, leite, iogurte e suco de frutas. Without the help of a dictionary, do you believe that you can understand at least ten items in this list and also the simple instructions in the note? (Reading-Basic 2)
15. Can you ask someone’s name in Portuguese? (OverallSpeaking-Basic 2)
16. You are invited to observe a trial in Brazil where a defendant is being accused of money laundering. Would you be able to understand all the details of the discussions? (Listening-Superior 2)
17. Of all the objects around you at this moment—e.g. ceiling, wall, chair, paper, etc, including your belongings (shoes, stocks, etc)—can you write a list of five of these objects, in Portuguese, without using a dictionary? (Writing-Basic 1)
18. Could you interpret Brazilian Portuguese for a person from your country in a casual setting—e.g. visiting a city—, for a period of one hour? (OverallSpeaking-Adv 1)
19. In your pronunciation can you distinguish words *pressente* and *presente, cassa* and *casa*? (PronunciationSpecifics-Basic 2)
20. Do you feel that you could shop for food or clothes in Brazil, speaking Portuguese, gesturing as needed, and make yourself understood? (OverallSpeaking-Interm 1)
21. Could you greet a Brazilian friend with simple expressions equivalent to the English expression “How are you?” in a casual setting? (OverallSpeaking-Basic 1)
22. Do you think that your pronunciation of Brazilian Portuguese is near-native?
23. If a Brazilian college student wrote to you a short letter of one page at most, in Portuguese, telling how he/she is doing, her accomplishments and plans, how his/her friends are doing, would you be able to read and understand the letter? You can use a dictionary. This can be any college student between 18 and 35 years old, who doesn’t use a lot of idioms or slang. (Reading-Interm 2)
24. With the help of a dictionary, would you be able to explain to your teacher in a note written in Brazilian Portuguese, that you had an appointment for a job interview, which conflicted with his/her class, and that was the reason for your being absent? It is ok to make a few mistakes. (Writing-Interm 2)
25. Would you be able to work as the editor of a newspaper written in Portuguese by college students at your school in the US? You would be required to write the editorials of this newspaper by yourself without serious grammar and cultural mistakes? (you may use a book, namely reference grammars, dictionaries, etc.) (Writing-Superior 1)
26. You are teaching first-year English to Brazilians in Brazil. One of the students who speaks only Brazilian Portuguese arrives late. He/She is uneasy about being tardy but tries to tell you what happened. Do you think that you would understand his/her full explanation of what happened? (Listening-Adv 2)
27. Suppose a Brazilian approaches you with some objects, probably to sell them to you, although you can’t tell for sure. The Brazilian notices that you are a foreigner and speaks slowly and with gestures. Would you be able to understand that the Brazilian wanted to sell the objects to you instead of something else? (Listening-Intermediate 1)
28. In Brazil, one often hears street vendors screaming repeatedly what they are selling. It can be ice cream, utensils, etc. Suppose a street vendor is passing by and selling several different products. Would you be able to remember and understand the names of five out of ten products a street vendor is saying if they are said in isolation, repeated 2–3 times? (Listening-Intermediate 2)
29. A Brazilian student will come to the US to stay in your house for one year. You are very happy to receive this Brazilian student. You have to write a two-page letter to him/her in Brazilian Portuguese, in which you try to explain about the US, what Americans are like, in your opinion. You also will need to talk about the activities to expect in the university/college environment in the US and what to wear, especially in the winter. It is ok to make a few mistakes. Would you be able to write the letter on your own, using a dictionary? (Writing-Adv 1)
30. If you had to look for a job in Brazil, would you feel comfortable making a phone call and setting an appointment for an interview, all in Portuguese? (OverallSpeaking Adv 2)
31. If a Brazilian you met said to you in Portuguese, “Hi! How are you?,” in a common, “television speaker” accent,

- without using idioms, do you believe that you would actually understand what he/she said? (Listening-Basic 1)
32. Have you read or studied from any source that the Portuguese language became the main language in Brazil only in the 2nd half of the 18th Century because the language Tupi-Guarani was the most common language spoken in Brazil until the 18th Century? (Cult/Civ-Superior 1)
33. If a Brazilian said to you in Portuguese, "What time is it?" in a common, "television speaker" accent, without using idioms, do you believe that you would actually understand? (Listening-Basic 2)
34. Would you be able to say briefly in Portuguese what you did last year? It would be all right to make a few mistakes, provided that you could communicate your ideas. Remember that speaking about events in the past requires the use of verbs in the past. (OverallSpeaking-Interm 2)
35. Suppose you have to write a two-page formal essay in Portuguese reviewing a short story you just read. This essay should include an introduction, a development and a conclusion. Could you write an essay with a few mistakes, using a dictionary? (Writing-Adv 2)
36. Did you know that Brazil is the only Portuguese-speaking country in Latin America? (Cult/Civ-Basic 1)
37. Suppose you were participating in a game. The purpose of the game is to let you choose five items from the list below so that you can drink and eat well. If you pick the wrong item you will be very sick. Would you be able to choose five items from the following list (there are more than five items that will be good for you): pão, carne estrogada, prego, água, lama, ovos, folha de bananeira, esterco, presunto, pó-de-serra, melancia, laranja, cenoura, leite, lavagem, farofa, pau, xícara, papelão, areia, banana, queijo, refresco, suco de frutas, carne de peru, frango, tijolo, pernil de porco, maçã, camarão e perfume. (Reading-Basic 1)
38. Suppose you are living with a host family in Brazil. Your Brazilian family leaves you a note explaining that they had to leave, where they went and why, and that they will be back late that evening. Would you understand the message with the help of a dictionary? (Reading-Interm 1)
39. The text below was created especially for this test. It contains grammar mistakes. Please study it carefully.
"En o filme Gladiador, Commodus é apaixonado por a própria irmã, Lucilla. Na verdade, Commodus tuvo relações incestuosas con várias de sus irmãs. Pervertido, mantenia um harém pessoal com cerca de 300 pessoas, entre mulheres e homens."
- Can you find and correct 7 mistakes or more. Repeated mistakes don't count. (Reading-Superior 1)
40. You need to write a list of groceries in Portuguese, so that the employee in a Brazilian supermarket will be able to bring all the groceries to your home. In your list you need to write your address, what you need, request that it be delivered to your home, and thank the person in a note of more or less two lines. You can use a dictionary to write your sentences and the words. Would you be able to write this message in Brazilian Portuguese? (Writing-Interm 1)
41. Do you think that you master the intonation and rhythm of Brazilian Portuguese like a native? (PronunciationSpecifics-Superior 2)
42. "Simpatia" is a common type of belief in Brazil that uses "recipes" to cure ailments, makes someone you love fall in love with you, etc. Have you heard of any of these recipes? (Cult/Civ-Interm 2)
43. The pronunciation of the r-sound in English is similar to the pronunciation of an "r" found in dialects of rural areas of Brazil, and generally not considered standard by teachers of Portuguese as a foreign language. In Brazil, this pronunciation is known as the "r-caipira". Do you think that you have mastered the pronunciation of "r" in Brazilian Portuguese, as it is generally pronounced in Brazil? (PronunciationSpecifics-Adv 1)
44. The text below is a passage of an actual popular song in Brazil, called Samba do Arnesto. It was written by Adorinan Barbosa and contains grammar that reflects the language variety from a low social class. The grammar in this song would not be acceptable according to standard grammar. Please study it carefully.

*O Arnesto nus convidô
 Prum samba, ele mora no Brais
 Nós fumu num encontremu ninguém
 Nós vortemu cuma baita дума reiva
 Da otra vez nós num vai mais
 Noutro dia encontremu co' Arnesto
 Qui pidiu disculpas mais nós num aceitemu
 Isso num si faziz Arnesto
 Nós num si importa
 Mais ocê divia ter punhado um recado na porta*

Would you be able to provide the correct form for 7 or more of the grammar mistakes? Repeated mistakes don't count. (Reading-Superior 2)

45. Have you read or studied any of the works of one of the following : Jorge Amado, Villa-Lobos, Machado de Assis, or Aleijadinho? If you just heard about them, you should answer "no." (Cult/Civ-Adv 1)
46. Imagine that you are in your house or apartment, in the room where you normally sleep. Would you be able, to write in Brazilian Portuguese, the name of any ten objects in your room in short, simple sentences that include these objects, without using a dictionary? These sentences can be of any kind provided that you vary the verbs (I like my books, I sleep in my bed, I have five pants, etc.). A few mistakes are allowed, but you cannot use a dictionary. (Writing-Basic 2)
47. In English, vowels are commonly lengthened or shortened in spontaneous conversation—e.g. the word "bonus" which has a long "o" and a short "u." Consequently, it is common for English speakers to mistakenly lengthen and shorten Brazilian Portuguese vowels when they speak Portuguese. Do you feel comfortable with your pronunciation of Brazilian

Portuguese vowels? (PronunciationSpecifics-Adv 2)

48. Some word pairs in Portuguese may be difficult for Americans to say. These are word pairs that have quite different meaning although their pronunciation seems similar to Portuguese students. These are words like *vi* and *vim*, *sal* and *são* etc. Do you think that you pronounce these words differently and correctly? (PronunciationSpecifics-Interm 2)

Table 1
Question difficulty levels

Question number	Difficulty	Question number	Difficulty	Question Number	Difficulty	Question	Difficulty
1	41.9	13	23.4	25	19.4	37	96.8
2	66.1	14	92.0	26	70.2	38	96.0
3	11.3	15	96.8	27	94.4	39	53.2
4	87.1	16	14.5	28	86.3	40	91.9
5	81.5	17	92.7	29	83.9	41	16.9
6	42.7	18	70.2	30	33.1	42	16.1
7	48.4	19	71.8	31	96.0	43	60.5
8	85.5	20	91.9	32	20.2	44	50.0
9	79.0	21	98.4	33	96.0	45	21.8
10	85.5	22	16.9	34	79.8	46	84.7
11	27.4	23	97.6	35	61.3	47	54.0
12	56.5	24	93.5	36	91.1	48	70.2

Table 2
Point Biserial item-total correlations

Question number	Difficulty	Question number	Difficulty	Question Number	Difficulty	Question	Difficulty
1	0.52 *	13	0.58*	25	0.42*	37	0.43*
2	0.27*	14	0.46*	26	0.54*	38	0.51*
3	0.46*	15	0.46*	27	0.44*	39	0.57*
4	0.62*	16	0.43*	28	0.55*	40	0.55*
5	0.55*	17	0.47*	29	0.52*	41	0.46*
6	0.52*	18	0.65*	30	0.62*	42	0.46*
7	0.49*	19	0.50*	31	0.42*	43	0.43*
8	0.33*	20	0.57*	32	0.18*	44	0.69*
9	0.22*	21	0.40*	33	0.45*	45	0.37*
10	0.37*	22	0.39*	34	0.59*	46	0.53*
11	0.48*	23	0.42*	35	0.63*	47	0.56*
12	0.34*	24	0.41*	36	0.21*	48	0.52*

Note: * significant at the 0.05 level

Table 3
Variation of the 48 items, separated in five groups through factorial analysis

Variance explained by each rotated factor

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Comprehension	Culture and civilization	Production (writing and speaking)	Miscellaneous	Pronunciation specifics
8.039%	5.718%	3.419%	7.234%	3.539%

Table 4

The five factors obtained through the factorial analysis and the language components in each factor. In parentheses are the number of items from each language area. These factors were labelled according to the highest number of language components shown in parentheses.

Factors obtained by the factor analysis Questions that loaded on each factor

Factor 1 Comprehension	Listening (5) – Reading (3) – Overall speaking (2) – Writing (1)
Factor 2 Culture and civilization	Culture/Civ (6) – Reading (1)
Factor 3 Production	Writing (7) – Overall Speaking (4) – Pronunciation Specifics (3) – Listening (1) – Reading (1)
Factor 4 Miscellaneous	Reading (3) – Culture/Civ (1) – Overall speaking (1)
Factor 5 Pronunciation specifics	Pronunciation specifics (5) – Listening (2) – Overall speaking (1) – Culture and civilization (1)

Table 5

Descriptive statistics of the subjective test

Statistics	Values
Mean	31.2
Median	32
Mode(s)	29, 33
Minimum	2
Maximum	48
Standard Deviation	8.3
Range	46
N=124 students	

Figure 1

Frequency distributions of BP self-assessment

