Using Video for Listening Assessment: Opinions of Test-Takers

David Progosh

This article reports on the researcher's use of video for listening comprehension in material design and test construction and discusses a survey questionnaire that solicited test-takers' views about video-mediated listening assessment in a tertiary institution in Tokyo. The results of the survey were to be used to justify (or not) the inclusion of video as a mode of presentation for listening assessment and to inform test designers of issues to consider for improving future video test design. It was found that the learners liked video in listening comprehension tests and preferred video tests over audiocassette tests. The findings will lead to further consideration and research in video-mediated assessment instruments.

Introduction

Vision is the most dominant perceptive sense humans use to process information (Gregory, 1972). In today's world, visual imaging systems have become extremely sophisticated and widespread. Video has become the most prominent medium for the depiction of moving images. Video is omnipresent in the industrialized (and not so industrialized) world, and it has become a regular part of our daily lives.

Educational planners recognize the usefulness of video as an instructional tool and have designed materials incorporating video technology for use in the classroom. In the field of teaching English as a second or other language, video has been closely associated with listening skills (Mueller, 1980; Ariew, 1987; Secules, Herron, & Tomasello, 1992; Baltova, 1995). Other ESL researchers have suggested video's effectiveness lies in its ability to motivate learner achievement (Oxford, Park-Oh, Ito, & Sumrall, 1993); to aid in understanding of paralinguistic aspects of communication (Lonergran, 1983; Stempleski & Tomalin, 1990); to provide a rich context for authentic discourse (Geddes & White, 1978; Tomalin, 1987; Altman, 1989); and to help solve problems in cross-cultural awareness and sensitivity (Kramsch, 1993; Stempleski & Tomalin, 1995).

The advocacy for use of authentic materials in the classroom (Cummins, 1989) also applies to video, and theoretical and practical frameworks have been suggested for using authentic video in the classroom (Ciccone, 1995). Lund (1990) produced a taxonomy of tasks for teaching listening skills and suggested task difficulty can vary while holding authentic texts constant across different levels of second-language proficiency. In short, there has

been much theory and research in using video as part of classroom pedagogy; however, discussion of video as a means for assessment is conspicuously absent.

In considering video as a medium for assessment the issue of economic and technological resources required for its implementation must be considered (Swain, personal communication, 1996). Compounding the issue is the differing perceptions of what the purpose and use of video are in listening assessment. Some test developers think of the visual channel as an aid to the audio channel (Pelletier, 1990; Dunkel, Henning, & Chaudron, 1990), whereas others perceive video as a medium with its own characteristics (Gruba, 1993, 1995; Benson & Benson, 1994; Kramsch, 1993).

Researchers have used quasi experimental techniques (van Lier, 1988) to control variables such as content, test items, and test design using different modes of presentation to make comparisons between test scores as a method to make claims about the effect of mode of presentation on language proficiency (Gruba, 1993; Baltova, 1995; Ryberg, 1995); however, Salomon and Clark (1977) showed comparisons between media were problematic because there is not enough control of the variables to draw valid conclusions.

Current psychometric definitions of listening for test purposes have neglected the influence visual elements have on our comprehension in everyday life. Test developers have, until now, steered clear of video-mediated tests, as construct validity is at risk (Bachman, 1990). Paradoxically, the validity of listening tests that do not take into account that most people both hear and see in most communicative situations is just as contentious.

Given the thorny issues, it is no wonder that researchers and test developers are wary of including video in assessment. Nevertheless, video is here to stay as a mode of presentation in the classroom, and the future promises even more use of video in areas such as satellite television, multimedia, and interactive video on computer networks such as the World Wide Web. It is timely that more investigations into the use of video in assessment be undertaken as this area is still in its infancy.

The Study

The study was carried out to determine test-takers' opinions of a video-mediated listening test used for assessing achievement in an EFL program that used video in the teaching of listening. Most of the research on test-takers looked at personal characteristics (e.g., background, motivation) that affect test performance (Zeidner & Benoussan, 1988; Bradshaw, 1990; Zeidner, 1990; Brown, 1993). Such research, despite suggestions that the findings can be useful in test design, have yet to be operationalized (Shohamy, 1982; Brown, 1993).

My survey instrument was designed to elicit the test-takers' views on a pilot video-mediated quiz of listening skills. The survey contained questions

about the level of difficulty of the test, the quality of the sound, the clarity of written instructions, the content of the test, and a comparison between video-mediated tests and the learners' previous testing experiences using audiocassette on achievement tests (see Appendix B).

It is important to note that the research is exploratory, as nothing in the literature has attempted to elicit the views of test-takers to video-mediated listening tests to date. Brown (1993) used a similar survey instrument to investigate test-takers' reactions to a *tape*-mediated test of *spoken* Japanese—clearly a different investigation. Buck (1994) reported on a study that used think-aloud protocols of what test-takers think about while doing a test of listening comprehension and concluded that the assumptions of unidimensionality in listening comprehension made by current norm-referenced psychometric models do not accurately reflect what test-takers actually do. Gruba (in progress) is using Buck's think-aloud protocol methodology to establish the nature of video-mediated listening comprehension.

The Setting
The research took place at a tertiary institution specializing in language and vocational studies in Japan. The school offered two program options: general English studies, with emphasis on English and language arts, and vocational English, for those students who wanted to specialize in a particular vocational area. All students had to take a 300-hour English program taught by native speakers of English. The mission statement of the institution was to give young (18-22 years old) Japanese students exposure to native speakers of English with the purpose of fostering a sense of internationalization through communication in English. To that end, a large faculty of native English teachers from around the world were employed. As primarily a two-year vocational institution, the school drew students who were interested in learning English for a variety of purposes: for future jobs, for travel, for study abroad, or as a hobby. Along with the English core program, there were courses in business English, grammar, and translation, taught primarily by Japanese instructors. The core English program was separated into reading and writing (RW) and speaking and listening (SL) skill classes. There were approximately 30 students in the RW classes and up to 15 in the SL classes.The vocational courses were of the business-secretarial nature, included typing, computer skills, and language arts, and were taught by specialists in each area.

Students were placed into one of three levels for core English instruction: high, intermediate, and low, based on scores on the reading and listening components of the Michigan placement test. High-level students' scores averaged from 425 to 800, intermediate-level students ranged from 250 to 424, and low-level students' scores were below 250.

DAVID PROGOSH

Role of the Researcher

As a supervisor for the intermediate program, it was my duty to make a course syllabus, design materials, and prepare quizzes and tests. An earlier course evaluation determined that the speaking/listening course was universally unappealing to both teachers and students because the materials were written for an ESL context in Britain, and much of the content had no relevance to Japanese learners in an EFL setting in Japan. Second, the audiocassettes that made up the listening component of the coursebook featured a standard British dialect that the students found "uninteresting." Third, non-British teachers, who made up the vast majority of the faculty, had difficulty in teaching the materials because they were not familiar with British formulaic and idiomatic expressions, nor the context of much of the material. Based on the results of the evaluation, I changed the syllabus, materials, and assessment procedures.

A Rationale for Video in Syllabus Design

A needs analysis was conducted and the results used as the basis for designing a course syllabus and for material selection. The needs analysis produced data on the types of topics students were interested in studying and student learning styles and strategies. Teachers' and administrators' views were also collected and considered in course design. The needs analysis showed that the students ranked watching movies as a highly motivating activity for the study of English, followed closely by interest in learning about foreign customs and international travel (Progosh, 1994).

Bearing in mind the mission statement of the institute, which was to provide Japanese students with exposure to native speakers of English, with the purpose of internationalization through communication in English, syllabus specifications were written so that they would be relevant to the needs of the students.

It was decided to incorporate video into the program for four main reasons: (a) the needs analysis showed the students were interested in studying from film and video; (b) a video production facility was available at the institute; (c) each classroom was equipped with a TV monitor and VCR; and (d) with a degree in film theory and production, I was interested in visual media.

Having a large expatriate faculty from a variety of countries was an invaluable resource for material design. Faculty members helped on the production of the videotape and listening tasks that would become the listening comprehension component of a speaking and listening textbook designed around topic areas the students indicated they were interested in studying. Faculty members were good-natured and willing to add their expertise to the video in various capacities, which lent a truly international flavor to the results.

A Rationale for Video Tests

Scholz (1993) argued that for learning, teaching, and testing to be in harmony, test specification and test items should be derived from the same needs analysis and program consensus used for developing instructional materials. Given that the listening comprehension component of the textbook was video-mediated, it followed that tests should also use video to satisfy construct and content validity (Bachman, 1990, Scholz, 1993).

Research in language testing has also pointed to the *washback* effect (Hughes, 1989; Shohamy & Wilson, 1992) testing has on learning and teaching. In addition to their use for making decisions about individuals, tests can be used to promote beneficial learning and teaching practices.

Designing the Video Test

A pilot test was designed based on the types of tasks developed for the coursebook (written specifically for the school), in which an entire unit devoted to movies highlights strategies in identifying movie genres, retelling the plot and theme, giving personal impressions, and talking about likes and dislikes. The video component of the unit depicts two people talking about various aspects of movies (acting, plot, theme, impressions) and their likes and dislikes about a popular Hollywood film. The students are to identify the speakers on the screen, answer multiple-choice questions about the dialogue, and indicate in a table what each person on the screen thinks about the film. A similar task was adapted for the test in this study (see Appendix A).

Another unit in the coursebook was about food. Again, this topic was highly rated by the students.[1] The video depicts a "cook" (an expatriate teacher) preparing a dish. The task has the students check off the ingredients used in the preparation of the dish and then write short answers based on the videotext. Again, a similar task was adapted for the test (see Appendix A).

As a first stage in test validation, it was decided to find out about the face validity of the test, that is, would the test be considered acceptable by teachers and students as a test? In order to obtain information on the perceptions of the test-takers to the test, a survey questionnaire was designed and administered with the pilot test.

Defining and Sampling the Population, and Research Method

The population was defined as all students in the second year of the intermediate program. It was not possible to administer the video-mediated pilot test and follow-up survey questionnaire to the entire second year intermediate population, so a sample was taken.

To obtain a sample that would be representative of the second year intermediate population, SL classes were chosen at random designated by class number code. The data indicate the sample was representative of the

intermediate population. The test was administered in the classes' regular classroom, which, as mentioned above, was equipped with a TV monitor and VCR. The survey questionnaire was administered immediately following the video quiz. The test-takers answered six closed-response questions; five used a seven-point Likert scale; and one was a binary choice on the survey (see Appendix B). It took approximately five minutes to complete. To analyze these data, the classes were collapsed into one group and analyzed as a whole (n=62) using the SPSS[x] program. Descriptive statistics (frequency, mean, standard deviation, minimum and maximum score) were calculated for each item.

Results

Question 1: How was the level of difficulty?
The test-takers were distributed normally in responding to this item, ranging from 2 (easy) to 6 (difficult), with a mean of 4.36 (SD=.986). This result indicates the test was perceived as being neither too difficult nor too easy, with the majority of scores clustered in the mid-range.

Question 2: How was the sound quality? Could you hear it clearly?
The mean is slightly negatively skewed suggesting the sound quality needed to be adjusted for the next video test (mean=4.468, SD=1.302). The result stimulated discussion about how to achieve better sound quality, and a number of factors were identified as possible problem areas, such as produc-tion techniques, quality of the TV monitors in the classroom, volume control, and the test-taker's position in the room. Subsequent video production produced better sound quality, and instructions for the classroom teacher were designed and made on how to set the volume of the TV monitor and where to seat test-takers for optimum viewing and listening clarity.

Question 3: Could you understand the written instructions?
The test-takers were required to read instructions on how to mark their answers on an answer sheet, and this question on the survey was included to measure the effect written instructions may have on test performance. Ideal-ly, test-takers would find the instructions easy to understand (a value of 1 on the Likert scale). The results showed the mean for this question was 3.290 (SD=1.151), indicating written instructions were not as easy as test designers thought they were, which affected test results. It was decided to include audio instructions on the videotape on how to answer the questions as part of future test construction.

Question 4: What about the content? Do you think it is a good idea to use movies for quizzes?
The results for this item were negatively skewed, with 92.2% of respondents falling into the *good idea* side of the median (mean=5.177, SD=1.443, range

1=bad idea to 7=good idea). This result was interpreted as an indication that face validity of the test met the approval of the test-takers, suggesting use of video in listening test design could be further explored.

Question 5: Compared to cassette tapes, is video easier or harder?
The distribution for this item is bimodal, which suggests the test-takers are split on their thoughts regarding whether video is easier or harder compared with audio, although the mean is slightly polarized to the *easier* side (mean=3.339, *SD*=1.342). This item was the most difficult to interpret, as factors that affect the test-takers' perceptions on why they responded the way they did were not solicited; the results indicate further research in test-takers' attitudes and comprehension strategies are well worth pursuing.

Question 6: Which method do you prefer: Audiocassette or video listening quizzes?
This item was a binary choice for test-takers, and 91.9 % of the respondents preferred video listening quizzes over audiocassette quizzes. Although there was no direct comparison of the two modes of presentation in the test situation, the test-takers were familiar with audiocassette modes of listening comprehension tests. The results were interpreted as indicating the video-mediated listening comprehension test satisfied the conditions of test construction suggested by Scholz (1993), Hughes (1989), and Shohamy and Wilson (1992), as interpreted and defined by the test designers in this specific setting (i.e., the test tasks satisfied construct and content validity).

Conclusion

The results of the survey are encouraging, as the sample think video in listening comprehension is a good idea, preferring video-mediated tests over audiocassette tests. Nevertheless, the purpose and use of the test must be kept in mind. As Scholz (1993) argued, for learning, teaching, and testing to be in harmony, test specification and test items should be derived from the same needs analysis and program consensus used for developing instructional materials. In this case, because the syllabus and materials incorporated video for listening comprehension, it followed that video should be used for assessing learner achievement. It has yet to be determined if such tests can be used for purposes of general language proficiency.

Further Research

Buck's (1994) research into the nature of listening comprehension presents us with a problem in current psychometric norm-referenced measurement models that assume unidimensionality in listening comprehension. He concludes that listening is not a unidimensional cognitive process and questions the validity of measurements that assume that it is. Clearly further research is needed in this area to expand our understanding of this complex process.

In the present study, the rather ambiguous results from question 5 indicate more research is needed into what makes the video-mediated test more difficult or easier for comprehension. Several areas of research need to be investigated: (a) what is the relationship of vision and hearing in overall comprehension of a videotext? (b) how does the comprehension process work? (c) what are the subjective differences in comprehending videotexts? (d) how do listeners monitor their own comprehension? (e) what effect do content, film genre, style, and technical values have on comprehension? (f) what is the role of cultural and background knowledge in comprehension of videotexts? and (g) what effect do test prompts and task type have on test response? Any of these areas is worthy of more in-depth exploration.

Note

[1]In Japan, a popular television program features celebrities cooking and tasting food from Japan and elsewhere.

The Author

David Progosh was a supervisor at Kanda Institute of Foreign Languages, a tertiary-level school in Tokyo, where he designed syllabi for EFL instruction for the past five years. He is co-author with Herman Bartelen of an ESL speaking and listening textbook titled *Opening Doors* (1995, Heinle & Heinle). He is currently pursuing studies in curriculum and assessment issues at the Ontario Institute for Studies in Education, Toronto, Canada

References

Altman, R. (1989). *The video connection: Integrating video into language teaching*. Boston, MA: Houghton Mifflin.

Ariew, R. (1987). Integrating video and CALL in the curriculum: The role of the ACTFL guidelines. In W.F. Smith (Ed.), *Modern media in foreign language education: Theory and implementation* (pp. 41-66). Lincolnwood, IL: National Textbook.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Baltova, I. (1995). The impact of video on the comprehension skills of core French students. *Canadian Modern Language Review, 50,* 507-523.

Benson, M.J., & Benson, V.A. (1994). The new challenge: Testing the video course. *IALL Journal of Language Learning Technologies, 27,* 11-26.

Bradshaw, J. (1990). Test-takers' reactions to a placement test. *Language Testing 7*(1), 13-30.

Brown, A. (1993). The role of test-taker feedback in the test development process: Test-takers' reactions to a tape-mediated test of proficiency in spoken Japanese. *Language Testing 10,* 277-303.

Buck, G. (1994). The appropriacy of psychometric measurement models for testing second language listening comprehension. *Language testing 11*(2), 145-170.

Ciccone. A. (1995). Teaching with authentic video: Theory and practice. In F.R. Eckman, D. Highland, P.W. Lee, J. Milham, & R. Rutkowski Weber (Eds.), *Second Language acquisition theory and pedagogy* (pp. 203-215). Mahwah, NJ: Erlbaum.

Cummins, P. (1989). Video and the French teacher. *French Review, 62,* 411-426.

Dunkel, P., Henning, G., & Chaudron, C. (1990). The assessment of an L2 listening comprehension construct: A tentative model for test specification and development. *Modern Language Journal, 77,* 180-191.

Geddes, M., & White. R. (1978). The use of semi-scripted simulated authentic speech in listening comprehension. *Audiovisual Language Journal 16*, 137-145.

Gregory, R.L. (1972). *Eye and brain*. New York: McGraw-Hill.

Gruba. P. (1993). A comparison study of audio and video in language testing. *JALT Journal, 16*(1), 85-88.

Gruba. P. (1995, March). *Design and development of video-mediated language tests*. Paper presented at the annual meeting of the Teachers of English to Speakers of Other Languages, Long Beach, CA.

Gruba, P. (in progress). *Towards a framework for the design and development of video-mediated foreign language listening comprehension tests*. Doctoral dissertation, University of Melbourne.

Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.

Kramsch, C. (1993). *Context and culture in language teaching*. Oxford: Oxford University Press.

Lonergran, J. (1983). Video applications in English language teaching. In J. McGovern (Ed.), *Video applications in English language teaching* (pp. 69-86). Oxford, Pergamon Press.

Lund, R.J. (1990). A taxonomy for teaching second language listening. *Foreign Language Annals, 23*(2), 105-115.

Mueller, G. (1980). Visual contextual clues and listening comprehension: An experiment. *Modern Language Journal, 64*, 335-340.

Oxford, R., Park-Oh, Y., Ito. S., & Sumrall, M. (1993). Learning a language by satellite television: What influences student achievement? *System, 21*(1), 31-48.

Pelletier, R.J. (1990). *Video in the evaluation process*. (ERIC Document Reproduction Service ED 316 022)

Progosh, D. (1994). Needs analysis: Topics. *Kanda Institute of Foreign Languages Academic Journal, 3*, 73-91.

Ryberg, K. (1995, March). *Video application research in ESL listening comprehension*. Paper presented at the annual meeting of the Teachers of English to Speakers of Other Languages, Long Beach, CA.

Salomon G., & Clark. R.E. (1977). Reexamining the methodology of research on media and technology in education. *Review of Educational Research, 47*(1), 99-120.

Scholz, G.E. (1993, April). *Exploring ESP and language testing*. Paper presented at the annual meeting of the Southeast Asian Ministers of Education Organization Regional Language Center Seminar, Singapore.

Secules, T., Herron, C., & Tomasello, M. (1992). The effect of video context on foreign language learning. *Modern Language Journal, 76*, 480-490.

Shohamy, E. (1982). Affective considerations in language testing. *Modern Language Journal 66*, 13-17.

Shohamy, E., & Wilson, A.R. (Eds.). (1992). *Language assessment for feedback: testing and other strategies*. Dubuque, IA: Kendall/Hunt.

Stempleski, S., & Tomalin, B. (1990). *Video in action*. New York: Prentice-Hall.

Stempleski, S., & Tomalin, B. (1995, March). *Using video to promote cultural awareness*. Paper presented at the annual meeting of the Teachers of English to Speakers of Other Languages, Long Beach, CA.

Tomalin, B. (1987). *Video, TV, and radio in the English class*. London: Macmillan.

van Lier. L. (1988). *The classroom and the language learner: Ethnography and second language classroom research*. London: Longman.

Zeidner, M. (1990). College students' reactions to key facets of classroom testing. *Assessment and Evaluation in Higher Education 15*, pp. 151-69.

Zeidner, M., & Benoussan, M. (1988). College students' attitudes towards written versus oral tests of English as a foreign language. *Language Testing 5*(2), 100-114.

Appendix A: Test Tasks

Task 1: Movies
Video

You will see two people talking about a movie. Before watching, read the questions. After you watch the first time, blacken the letter of the best answer.

1. What is the name of the movie?
a. Terminator 2.	b. Terminator.
c. Terrapin Station.	d. Time Cop.

2. What type of movie is it?
a. science fiction.	b. far fetched.
c. action.	d. love story.

3. Does the man on the right like the acting?
a. yes.	b. no.

4. What does the man on the left dislike about the movie?
a. the acting.	b. the sets.
c. the cinematography.	d. the special effects.

5. Does the man on the right think the movie is expensive to make?
a. yes.	b. no.

6. What scene do they talk about?
a. a scene when a car crashes.
b. a scene when the actor cuts his elbow.
c. a scene when the actor cuts his eye out.
d. a scene when the actor is killed.

7. What do they agree about?
a. the acting.	b. the sets.
c. the cinematography.	d. the special effects.

What kind of movies do the men like? After you watch the second time, put a check mark (4) in the box they mention.

	Type Of Movie	*Man On The Left*	*Man On The Right*
8.	action		
9.	Hollywood		
10.	foreign		
11.	documentary		
12.	love story		

Task 2: Food
Cooking With Mr. Pease
Video

Watch and listen to Mr. Pease and his helpers make guacamole, a well-known Mexican appetizer.

1st Viewing

While watching the video, put a check mark (4) beside the ingredients used.

avocado	eggplant	salt	butter
tomatoes	pepper	potatoes	onion
garlic	lime	chili sauce	chili pepper
sour cream	sugar	lemon	vinegar

2nd Viewing

Write down the reasons why Mr. Pease says the following statements.

| Statement | Reason |

1. It's one of the easiest dishes in the world to make.
2. You must eat it right away.
3. Put them on the refrigerator for a few days because ...
4. Be careful with the chili sauce.
5. Don't forget the chips.
6. If you order "avocado" in Mexico, you have to be careful because they might bring you a lawyer.

Appendix B
The Survey Questionnaire

Survey: Video Listening Mini Quiz

Please answer the questions about the Video quiz. Circle your answer.

1. *How was the level of difficulty?*

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Very Easy   *So-So*   *Difficult*

2. *How was the sound quality? Could you hear it clearly.*

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Very clear   *So-So*   *Not clear*

3. *Could you understand the written instructions?*

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Very easily   *So-So*   *Difficult*

4. *What about the content? Do you think it is good to use movies for quizzes?*

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Bad idea   *So-So*   *Good Idea*

5. *Compared to cassette tapes, is video easier or harder?*

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

Easier than cassette   *Harder than cassette*

6. *Which method do you prefer: audio-cassette or video listening quizzes? Circle your choice.*

     Video Quizzes   *Audio-Cassette quizzes*