

# Teachers and learners evaluating course tasks together

Timothy Stewart

*This paper describes an approach to task evaluation that emerged out of the process of the negotiated development of a course between two co-teachers. The course was co-taught by one ELT specialist and a specialist in a subject area. The teachers were equal partners in this self-contained course. While teaching a new class, the teaching partners sought students' opinions on course tasks. They acknowledged that with classroom communication largely in teacher control, students often struggle to understand the patterns of communication presented. The assumption is that this can result in different interpretations of and participation in classroom activities by students. In the interest of generating multiple observations and evaluations of tasks, the co-teachers created a multi-layered reflection process. This process synthesizes student and teacher assessments of tasks written in learning logs with more traditional course evaluation data, in a reflective process of course development.*

## Introduction

This paper describes an approach to task evaluation designed to try and narrow the gap between teacher and learner perceptions of learning tasks. Teachers often talk about what worked in lessons, but most do not know much about what their learners think about the tasks they use. Much of what teachers do know is either through summative evaluations, or intuitive reflection (Burns 1999; Genesee and Upshur 1996).

At an English-medium liberal arts university in Japan, I team-taught an integrated language-content course on Issues in Cross-cultural Communication with a specialist in that field. Credit courses in the first and second years at the university are team taught by pairs of language and discipline-area faculty. In the second term of their second year, all students participate in a study-abroad programme. Our class employed task-based learning in order to prepare Japanese university students for a semester of studying abroad in English speaking countries. The 20 second-year students in the class ranged in speaking/listening proficiency from low- to high-intermediate level.

Teachers 'tend to assume that the way we look at a task will be the way learners look at it. However, there is evidence that while we as teachers are focusing on one thing, learners are focusing on something else' (Nunan 1989: 20). My partner and I wondered: how could we find out what our learners think about course tasks, and how might students' perceptions of

tasks converge with and diverge from our own? Our goal was to use this information to improve future editions of the course.

Learners transform a task through their reinterpretation of it (Breen 1989) and their interpretation determines actual learning outcomes (Littlejohn and Windeatt 1989). I argue in this paper for a multi-faceted approach to task evaluation as a way to match teacher and learner impressions of task appropriateness, in order to gain a clearer picture of the learning generated by tasks. Thus, I question, along with others such as Bailey *et al.* (2001), how much teachers can know about the appropriateness of classroom tasks without a multi-layered evaluation.

## Classroom-based evaluation

A useful model for classroom-based evaluation is outlined by Genesee and Upshur (1996), who delineate four essential components: having a purpose for evaluation, collecting information, interpreting information, and decision-making. They stress that, 'decisions are based on informed judgment . . . [and] require the careful collection of relevant information and a thoughtful interpretation of that information' (*ibid.*: 4).

Classroom-based evaluation concerns 'taking action to reduce mismatches' (Genesee and Upshur 1996: 40). The multi-layered process to evaluation described here represents one way for promoting teachers' understanding of learner perceptions. Each layer adds a deeper level of understanding. Thus, it is the inverse of the peeled onion metaphor. As new information is collected and interpreted, reflection on the data is essential.

The class was co-taught by me, an ELT specialist, together with a professor of cross-cultural studies. We designed the course tasks, headed by two research projects, with the study-abroad requirements for students at our university in mind. We were equal partners in the course and as such jointly created materials, taught, and determined grades. The course was self-contained and we worked simultaneously in the classroom. Through this collaboration, a model for course/task assessment took shape.

We decided to evaluate the tasks with learning log journals (Genesee and Upshur 1996). These journals were structured to focus students' and teachers' entries on evaluating specific course tasks. (See Appendix.) Two teachers and 20 students wrote learning log entries over the period of one 15-week teaching term. Each of the eight log entries contained perceptions of learners and teachers on the learning acquired as a result of doing specific tasks.

Tasks evaluated were mostly longer sequences of instruction extending over several lessons, linked by themes. These tasks varied in length from one lesson to several lessons. They were primarily concerned with meaning; related to the world outside the classroom; focused on task completion; and assessed in terms of task outcome. This follows the definition proposed by Skehan (1998) who contends: 'What counts, in task-based approaches, is the way meaning is brought into prominence by the emphasis on goals and activities' (*ibid.*: 268). In other words, tasks need outcomes to motivate learners into participation. We flagged eight tasks for evaluation. (See Table 1.)

| Tasks  | Average Ratings*            |
|--|-----------------------------|
| <b>1 Shaka-Dagang simulation</b><br>Adaptation of cross-cultural simulation 'BaFa BaFa' designed by R. Shirts. Simulated experience of interacting with a new culture.   | S = 3.5<br>Ta = 4, Tb = 4   |
| <b>2 What is culture?</b><br>Students first brainstormed in groups on the question: 'What is culture?' Each student then wrote their own definition of culture. Next, five elements of culture (behaviour/practice, material culture, norms, values and worldview) were taught through a series of worksheets and activities. This led to negotiations on new definitions of 'culture'.                                      | S = 3.2<br>Ta = 3, Tb = 4   |
| <b>3 Model research project</b><br>To compare elements of traditional/contemporary Japanese culture, students went individually to either a franchised or a non-franchised restaurant and entered data onto worksheets. Data later compiled and compared. Data analysed in a series of steps leading to a final presentation that was meant to conclude with comments on Japanese culture as seen through these restaurants. | S = 3.3<br>Ta = 3-, Tb = 3  |
| <b>4 Core concepts review</b><br>Basic review of course content to date.   | S = 3.5<br>Ta = 4, Tb = 3   |
| <b>5 Final research project</b><br>Modelled on 3 above, but designed by learners with guidance.  | S = 3.3<br>Ta = 3, Tb = 3.5 |
| <b>6 PowerPoint presentations</b><br>Final research project paper made into slide-show presentation.   | S = 3.5<br>Ta = 3, Tb = 3.5 |
| <b>7 Watch and report news presentations</b><br>Individual student presentations of current news.  | S = 3.7<br>Ta = 4, Tb = 4   |
| <b>8 Barnga simulation</b><br>Students learnt a card game. Specific rules of the game were slightly different at each table. This created some tension and misunderstanding and simulated cross-cultural experiences. Debriefing activities followed.  | S = 3.6<br>Ta = 3, Tb = 4   |

TABLE 1  
Evaluated course tasks

### Multi-layered approach to task evaluation

Note: Tasks appear in the order in which they were evaluated. \*Ratings on a scale of 1 to 4, where 4 is most effective.

Teachers conducting classroom-based evaluation want procedures that are practical. Information generated needs to be relevant to the purpose and situation. Finally, teacher-driven evaluation has to be useful for making decisions. According to Genesee and Upshur (1996: 6): 'One might undertake evaluation in order to make decisions about follow-up instruction for an entire class or to ascertain the effectiveness of particular

instructional units with a view to improving them'. A multi-layered approach can be used for either, or both. Our objective for this evaluation was the latter.

The approach has evolved into the layers shown in Table 2. It began simply as just the learning logs, layers 1 and 2. We wrote learning logs in class at the end of a selected task. Layer 3 was added since we needed to code all student entries and summarize the teacher evaluations after the course ended. What we sought were 'recurring patterns or salient events' (Bailey 1990: 215), but we recorded all idiosyncratic comments too. Once we had the learning log data coded, we started to think more carefully about how to interpret it. This led to the next stage of reflection which compared the teacher evaluations to the coded student evaluation summaries. We decided to first write individual reflections on the data from layers 1–3. We then read each other's reflections before having a discussion. The Layer 4 discussion phase is where we made tentative decisions about the course. After this, we tabulated the course-end questionnaire data to compare with the journal evaluations of tasks. Finally, we discussed possible meanings of the information and made decisions on course development. Since learning log reflections were written during lessons, the entire approach took approximately eight hours outside of class. For us, this was not much 'extra' time to spend on course development since we generated data upon which to make informed decisions.

| Layers                           | Completion time         |
|----------------------------------|-------------------------|
| 1 in-class teacher learning logs | 10 minutes during class |
| 2 in-class student learning logs | 10 minutes during class |
| 3 summaries of log entries       | 3 hours                 |
| 4 teacher reflections on 1–3     | 2 hours                 |
| 5 course evaluation survey data  | 1 hour                  |
| 6 course development decisions   | 2 hours                 |

TABLE 2  
Multi-layered approach  
to task evaluation

A major aim of this multi-layered approach is to employ multiple observers and multiple evaluations of tasks in order to gain a prism-like view of learning tasks. Because both teachers, and all students in the course, responded to the same set of questions, a multi-perspectival effect was created. Later, an end-of-course questionnaire data was added. Thus, by generating rich, multi-layered, and multi-perspectival data, my partner and I were able to reflect upon differing perspectives both in writing and discussion. To avoid confounding effects, we did not read other logs until after writing our separate entries.

The crucial stage of analysis in Layer 3 followed an inductive approach in which the categories that emerged reflected directly on the participants' log entries. Once the data were coded according to salient categories, both instructors reviewed all entries to assure coding accuracy.

The multi-layered  
process at work

Using the following example, I endeavour to guide readers through the multi-layered approach to task evaluation (Table 2) in something of the way

I experienced it as a teacher. This example actually covered several tasks up to the mid-term point. The initial task was an exploration of the meaning of the term 'culture'. Our teaching objective was to introduce concepts (i.e. behaviour/practices, material culture, norms, values, and worldview) that could be applied throughout the course. To start, the students practised writing extended definitions that helped them to produce their own definition of culture. Next, we used various materials to teach the key concepts listed above. Task 2 took four class periods (10 hours) to complete and ended with group negotiations in English on a definition of culture.

### Layer 1: Teacher perceptions

I begin with my own initial perceptions. My evaluation of Task 2 shows that I discovered how difficult it was for our students to comprehend abstract concepts. The results were that the learning of these concepts extended throughout the course, during which my impressions of this sociological approach to the course shifted. I rated Task 2 as 3 out of 4 based largely on quiz scores that showed learners were all able to identify the concepts. In the multiple-choice section, only four students did not get a perfect score. I saw this as indicative of achieving a basic level of understanding of the concepts taught. However, students continued to struggle with the concepts.

Name: Teacher A      Date: April 26, 2001      Task 2: What is culture?

Journal entry #2 – excerpt [see Appendix]

(2) Our main objective was to teach the parts of culture along with language to talk and write about these ideas. The unit concluded where it began with a definition of culture. Students considered examples in groups so that they could see the concepts from many angles to gain understanding. We discovered that this was not easy and acknowledged the difficulty of the task and of the concepts themselves. However I believe that most of the students now have a deeper understanding of what underlies human behaviour and are able to ask questions about it and explain some aspects of Japanese culture using the analytical framework presented.

(3) I think that much (3) learning was generated by this task.

(4) The quiz (77% average) showed that students were able to write something intelligible about these concepts... some students are struggling with the language of definition and with a clear knowledge of the concepts, however. We have much work to do both on the language and the content sides to reinforce what the students have learnt so far.

FIGURE 1  
Sample teacher learning  
log entry

We reviewed this initial framework for analysing culture after introducing more core concepts. I rated this mid-term review a top score of 4. The reason was that 12 of 20 students achieved 'A' grades whilst

just two students failed the mid-term test. I did note, however, that the test material was ‘thoroughly reviewed in lessons and graded liberally’. At the mid-term point, I was certainly favouring this approach to the course.

My confidence wavered during the two research projects. Conclusions presented in the model research project and the final research project (Table 1) signalled to me that many learners were not applying the main theoretical constructs to their analysis. As a result, many conclusions were terribly superficial. This called into question the appropriateness of the sociological orientation chosen for the course. Language did not seem to be a major issue as we provided learners with much support so that they were able to express themselves reasonably well. The cognitive load of the conceptual framework appeared to overwhelm many of them. In content-based courses this is a common concern. My evaluation of the PowerPoint Presentation task revealed cautious optimism. I rated the task a 3, but noted some presenters did not make clear connections between the data collected and their hypotheses and conclusions. Some failed to include any conclusions at all. Remember, we wrote entries separately and did not read other evaluations until the course ended.

My partner’s evaluations indicated that she held strong reservations about the transparency of these tasks for our students. Her concerns were: our initial objectives were too ambitious, the concepts were too abstract, our directions and presentation were not clear enough to make objectives transparent to students, and students continued having great difficulty interpreting data. I felt a sense of relief because she had opted for this sociological approach to the course so I was concerned that perhaps she felt some deep investment in it. Next, I needed to see what the students thought before we talked about the future shape of the course.

### **Layer 2: Student perceptions**

In class, we introduced the learning logs to our students just after concluding Task 1. Every student was given a numbered B5-sized notebook. We explained to the students that our aim was to try and learn what they thought about the activities we used in the course. We told them that we would *not* read their logs until after the course concluded. We presented two models evaluating one simple task completed in the previous lesson to illustrate how students could approach their learning log entries. Each model was written by one of the instructors and discussed in detail just before we all made our first learning log entry. These were available to learners as OHP transparencies whenever they evaluated lesson tasks. All student and teacher learning log entries were written during lessons. Logs were collected for safekeeping by the teachers immediately after each entry was completed. The students were told that their participation was optional.

Task 2: What is culture? \* 4/26/2001

(1) Were you absent for part of this task?

No.

(2) What did you learn from this task?

I learnt what is behaviour, material culture, norms, values and world view. It was very difficult to understand for me, but I tried to understand it.

(3) How much did you learn from this task?

I learnt very much (4).

(4) Explain your reasons for the above rating.

I didn't know anything about kinds of culture so I learnt a lot of new things in this activity.

\* Reproduced excerpts of student #6 learning log. Selected at random.

FIGURE 2  
Sample in-class student learning log entry

### Layer 3: Learning log summaries

The learning logs contained raw data for analysis. At the end of the course my co-teacher and I read through all learning logs, and created categories out of the data that generalized ideas found in the entries. This analysis allowed us to see the number of points raised by students, the number of students who mentioned a particular point, and which students concurred on a point. Thus, we could identify what the learners saw as strengths and weaknesses of particular tasks. Without this step in the process, it would be impossible to determine the patterns of responses and the issues needing reflection.

| What students said they learnt<br>(edited to top three)  | Log Number                               |
|--|--|
| 1 I learnt the elements of culture (reiterated concepts taught).                                   | 1, 2, 5, 6, 12,<br>13, 16, 17,<br>18, 20 |
| 2 I realized the depth of culture and the extent to which it influences our behaviour and beliefs. | 2, 6, 7, 11,<br>14, 18, 20               |
| 3 I think it's useful/important to know (no details).  | 1, 3, 4, 9, 15                           |
| Average rating = 3.2 (rated on 4.0 scale)  |  |

FIGURE 3  
Student learning logs summary sample (What is culture? (26 April 26 2001))

The students raised serious concerns about the theoretical orientation in their reflective evaluations. This was true in the evaluations of all five tasks related to the theoretical concepts. Four students said that the tasks and/or concepts were difficult to understand, and that their purpose and relation to

the course was unclear. Student entries about Task 2 that introduced the key concepts were somewhat empty as many stated simply that 'it's useful' or 'important to know' with no supporting details. By far the majority of comments were just a listing of the elements of culture that we taught. The same was true on the evaluation of the tasks we did to review the key concepts at mid-term. Whilst, several students mentioned language and skills they had learnt, the vast majority simply said that it was 'good' to review the concepts, possibly because they did not comprehend them well. Even though the students rated tasks 2 and 4 highly at 3.2 and 3.5 respectively, their comments were reserved.

#### **Layer 4: Teacher reflections on log data**

After reviewing the tabulated learning log data, my partner and I independently wrote summaries of what we saw as the main issues for discussion. We then read both of these summaries in preparation for our first formal discussion on course development, in which we tentatively set changes to the course for the following year.

##### **Teacher A:**

For me, the most striking contrast between teacher/student reflections was that we, as the course instructors, were nearly exclusively focused on the processes we hoped that the students would learn and the language teaching opportunities, whereas the students were overwhelmingly concerned with issues of content learning. Several of our entries indicated our perception that the students were not associating the theoretical concepts taught with the field research projects that they were engaged in. Yet, I believe now that we were wrong. In fact, there is much evidence to suggest the opposite. Many students distinctly wrote that they learnt about Japanese society and culture change generally. Others stressed becoming more aware of their own actions and unconscious assumptions. Since these and other connections did not seem to be apparent to us in their course work, we might conclude that the students did not have enough opportunity to fully process this new information to use it actively for interpreting data and situations. Perhaps we need to simplify the course tasks to give them more opportunities to apply theoretical concepts.

##### **Teacher B:**

Systematically analysing students' log entries made me realize that in my attempt to teach students research methods and the academic English language that goes with this, I lost sight of the students' learning about the content of their research projects. While we were focusing on how students analyse and synthesize data that they gather, students were focusing on what they were learning as a result of doing the research. They had difficulty applying the theories to the projects. I could have made the connections between the concepts and the projects more transparent.

**FIGURE 4**  
Sample of teacher reflections on log data (What is culture? (October 2001))



### Layer 5: Course evaluation data

Our end-of-course questionnaire is an anonymous survey that asks students for a 1 to 4 rating on each task listed in the categories of enjoyment, English learning, and content learning. At the end of a course, students may be better able to see the purpose of tasks and how sets of tasks link. This kind of data is very useful in a supplementary role, but summative surveys are no substitute for data collected immediately after tasks. In the end-of-course survey, 'Theories and concepts about culture' was rated as the third lowest task for enjoyment, English and content learning. It is apparent that our students saw learning value, both in terms of language and content, in the tasks that built the theoretical base, yet did not 'enjoy' them much.

TABLE 3  
Sample end-of-course  
questionnaire data

| Task   | Enjoy | English | Content |
|--|-------|---------|---------|
| Theories and concepts about culture<br>(4.0 scale) | 2.7   | 3.1     | 3.3     |

### Layer 6: Course development decisions

In this final stage, we discussed the information generated and moved to act. We decided to change the theoretical framework of the course and to reduce the theoretical structure significantly. We also streamlined the Model Research Project, trying to give it a sharper focus. By de-emphasizing the theoretical framework, we intended to bring more accessible experiential material to the fore. Hence, we chose to drop the sociological slant in favour of a more psychological approach.

## Conclusions

### Practicality

The approach can be used to evaluate one or several course tasks. Evaluations are written during lessons and the most time-consuming part, analysing learning log content, can be done after a course ends. It took us roughly eight hours of time outside of lessons to evaluate a full set of course tasks. A written record is produced that creates a dynamic understanding about task appropriateness and facilitates decision-making. It is likely that teachers would only want to use this style of evaluation when developing new courses and tasks.

### Relevance

Both qualitative and quantitative data are produced. The in-class student and teacher task evaluations ask participants to rate the learning generated by tasks on a four-point scale. This is also done for the end-of-course questionnaires in separate categories. Other information is in the form of qualitative comments that are later quantified as they are coded. Information collected in the reflective learning logs was *fresh* because participants were asked to evaluate tasks immediately following their completion. Reflective evaluation by *all* participants in a course is the main strength of the comparative approach described in this paper. Contrary to studies that feature learner diaries written by linguists (McDonough 2002), this is a case of regular students and their teachers recording impressions of learning. Unique to this approach is that teachers and learners evaluate tasks simultaneously. If learning diaries are kept only by teachers, their viewpoint gets reinforced (McDonough 1994), thus serving to perpetuate teacher-centredness. One potential weakness is that we must trust the sincerity and accuracy of entries. Also, entries may be fuller when written in the learners' native language. A further weakness of the approach is that it is

not directly used for on-line course development. Therefore, it does not respond to student concerns until the following year.

## Usefulness

My partner and I shared a sense that our journal reflections were obvious. Furthermore, we talked at length about how, after initially scanning the student logs, we felt their entries did not contain much to reflect upon. Later, like others (Barkhuizen 1998; Block 1996), we were surprised by what we learnt when we analysed and compared data. The multiplicity of perspective gained from observing the same phenomenon is one of the main benefits of collaborative evaluation and reflection.

Because of the influence of learners on lesson outcomes (Breen 1989), task-based pedagogy success should be measured through the 'degree to which teacher intentions and learner interpretation of a given task converge' (Kumaravadivelu 1991: 100). Concern for closing the gap in perceptions to facilitate decision-making evolved the multi-layered approach to task evaluation. While teachers' knowledge cannot be discounted, it encompasses only one angle on the complex picture of classroom interaction. With the multiple perspectives offered by the approach described above, a more holistic picture of learning can be built. Having access to varied interpretations of tasks can provide a clarifying effect. We felt empowered to take action to improve our course and saw ways to do so.

*Revised version received June 2005*

## References

- Bailey, K. M.** 1990. 'The use of diary studies in teacher education programs' in J. C. Richards D. Nunan (eds.). *Second Language Teacher Education*. New York: Cambridge University Press.
- Bailey, K. M., D. Freeman, and A. Curtis.** 2001. 'Goals-based evaluation procedures: how students perceive what teachers intend'. *TESOL Journal* 10/4: 5-9.
- Barkhuizen, G. P.** 1998. 'Discovering learners' perceptions of ESL classroom teaching/learning activities in a South African context'. *TESOL Quarterly* 32/1: 85-108.
- Block, D.** 1996. 'A window on the classroom: classroom events viewed from different angles' in K. M. Bailey and D. Nunan (eds.). *Voices from the Language Classroom: Qualitative Research in Second Language Education*. Cambridge, UK: Cambridge University Press.
- Breen, M.** 1989. 'The evaluation cycle for language learning' in R. K. Johnson (ed.). *The Second Language Curriculum*. Cambridge: Cambridge University Press.
- Burns, A.** 1999. *Collaborative Action Research for English Language Teachers*. Cambridge, UK: Cambridge University Press.
- Genesee, F. and J. A. Upshur.** 1996. *Classroom-based Evaluation in Second Language Education*. New York: Cambridge University Press.
- Kumaravadivelu, B.** 1991. Language-learning tasks: teacher intention and learner interpretation. *ELT Journal* 45/2: 98-107.
- Littlejohn, A. and S. Windeatt.** 1989. 'Beyond language learning: perspectives on materials design' in R. K. Johnson (ed.). *The Second Language Curriculum*. Cambridge: Cambridge University Press.
- McDonough, J.** 1994. 'A teacher looks at teachers' diaries'. *ELT Journal* 48/1: 57-65.
- McDonough, J.** 2002. 'The teacher as language learner: worlds of difference?' *ELT Journal* 56/4: 404-11.
- Nunan, D.** 1989. *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.
- Skehan, P.** 1998. 'Task-based instruction'. *Annual Review of Applied Linguistics* 18: 268-86.

## The author

**Tim Stewart** has team taught courses with ELT and subject-area specialists for over 15 years. He is a founding faculty member of Miyazaki International College in southwestern Japan. After

reflecting on his practice with colleagues for 10 years there, he accepted a tenured position at a public university in Japan. He is currently helping to

develop a new Department of Communication and Information Studies at Kumamoto University.  
**Email: [stewart@kumamoto-u.ac.jp](mailto:stewart@kumamoto-u.ac.jp)**

**Appendix: Learning log prompts**

Date: \_\_\_\_\_

Task name: \_\_\_\_\_

1 Were you absent for part of this task?

YES

NO

2 What did you learn from this task?

3 How much did you learn from this task? (Choose one)

very little

very much

1 2 3 4

4 Explain your reasons for the above rating.