Conducting Research on Performance-Assisted Learning

David Kluge

Nanzan University

klugenanzan@gmail.com

Abstract

Teachers who include performance activities in their course syllabus may find it difficult to convince administrators and other teachers of the value of its use, and one of the reasons is a lack of research data on the efficacy of performance use in teaching. This article proposes that practitioners of Performance-Assisted Learning (PAL) conduct micro-evaluations of the PAL activities that they do in class. The aggregation of such evaluations will form a macro-evaluation of PAL and may provide support for the use of PAL. The micro-evaluations would be based on the Ellis (1997) article on task evaluation which is described in detail in this article.

Barbee (2016), after listing eleven verbal, cognitive, affective, social, and educational benefits of a performance activity from Maley and Duff (1978, 2011), went on to claim that his implementation of the activity was prohibited by the program administrator at his university. Barbee reported: "I was told directly that dramatic activities were unproductive and did not have the *appearance* of being academic enough; I was told directly not to 'play games'" (p. 6). Carpenter (2015) reports the same problem with implementing drama. Another problem with implementing performance such as drama, debate, and oral interpretation in the language class is the perception on the part of teachers that special expertise is required (Kawakami, 2012; Kluge & Catanzariti, 2013). Much of this hesitance to adopt performance activities is due to a lack of understanding about the role of performance in learning, but also to a lack of research on the benefits of such activities. This research deficit and how to ameliorate it is addressed in this article. This article first defines the blanket term for using performance in learning, Performance-Assisted Learning (PAL), then explains how PAL activities can be

evaluated through research using a model by Ellis (1997), and finally outlines the proposed micro-evaluations that will lead to a macro-evaluation of PAL.

Performance-Assisted Learning Definition

Performance-Assisted Learning (PAL) is "using any kind of performance to assist in the learning, consolidation, and assessment of content" (Kluge in Head et al., 2018, p. 234). It is distinguished from Performance-Based Learning where the performance is the main part of the course, whereas in PAL the performance is a task that is used within a typical curriculum or syllabus. In other words, PAL is a Task-Based Learning (TBL) activity, whereas Performance-Based Learning is a Project-Based Learning (PBL) activity.

Ellis' Model for Task Evaluation

In a short seven-page article on textbook evaluation, Ellis (1997) states that teachers can evaluate their textbooks using one of two ways: "impressionistically or they can attempt to collect information in a more systematic manner (i.e. conduct an empirical evaluation)" (p. 37). He suggests that one way to evaluate a textbook empirically is through evaluating the tasks contained in the textbook. Ellis then describes how to conduct such a task evaluation. He first distinguishes between macro-evaluation and micro-evaluation. About macro-evaluation, Ellis (1997, p. 37) states, "A macro-evaluation calls for an overall assessment of whether an entire set of materials has worked. To plan and collect the necessary information for such an empirical evaluation is a daunting prospect." He goes on to describe micro-evaluation (Ellis, 1997, p. 37): "In a micro-evaluation, however, the teacher selects one particular teaching task in which he or she has a special interest and submits this to a detailed empirical evaluation." Most importantly, he concludes, "A series of micro-evaluations can provide the basis for a subsequent macro-evaluation" (Ellis, 1997, p. 37). This conclusion forms the basis for the proposed macro-evaluation of PAL through a large collection of micro-evaluations of a variety of PAL tasks.

Task Definition

Ellis (1997, p. 38) describes how to conduct a micro-evaluation of tasks. He first defines "task" (emphasis added):

> This term is now widely used in language teaching methodology (e.g. Prabhu 1987; Nunan 1989), often with very different meanings. Following Skehan (1996), **a task is here viewed as 'an activity in which meaning is primary; there is some sort of relationship to the real world; task completion has some priority; and the assessment of task performance is in terms of task outcome'**.

There are four components in the Ellis/Skehan definition. Many PAL activities can easily be seen to contain each component:

1. **"a task is here viewed as 'an activity in which meaning is primary"**
   In PAL activities (e.g., speeches, presentations, dramas, roleplays, debates, etc.), communication of meaning is paramount as the role of the performer is to communicate meaning to an audience.
2. **"there is some sort of relationship to the real world"**
   Most PAL activities mimic real-world activities, especially drama and roleplays, but also speeches, presentations, and debates are done in some form in the workplace or in real-world organizations such as government and business.
3. **"task completion has some priority"**
   Most PAL activities such as the ones mentioned above are required to have a clear beginning, middle, and end, so task completion is a necessary PAL component.
4. "**the assessment of task performance is in terms of task outcome"**
   Since many PAL activities are meant to be performed in front of an audience, the performers are judged on how well they completed their task.

Skehan's definition is the one that will be used for this paper as many PAL activities clearly fit as tasks in the Skehan and Ellis sense.

Evaluating a Task

Ellis (1997, p. 38) continues by listing the seven steps to evaluating a task in four stages, with Steps 1-3 comprising the Preparation stage of the evaluation, Step 4 as the Implementation stage of the evaluation, Steps 5-6 as the Evaluating stage, and Step 7 as the Reporting stage of the evaluation (see Figure 1):

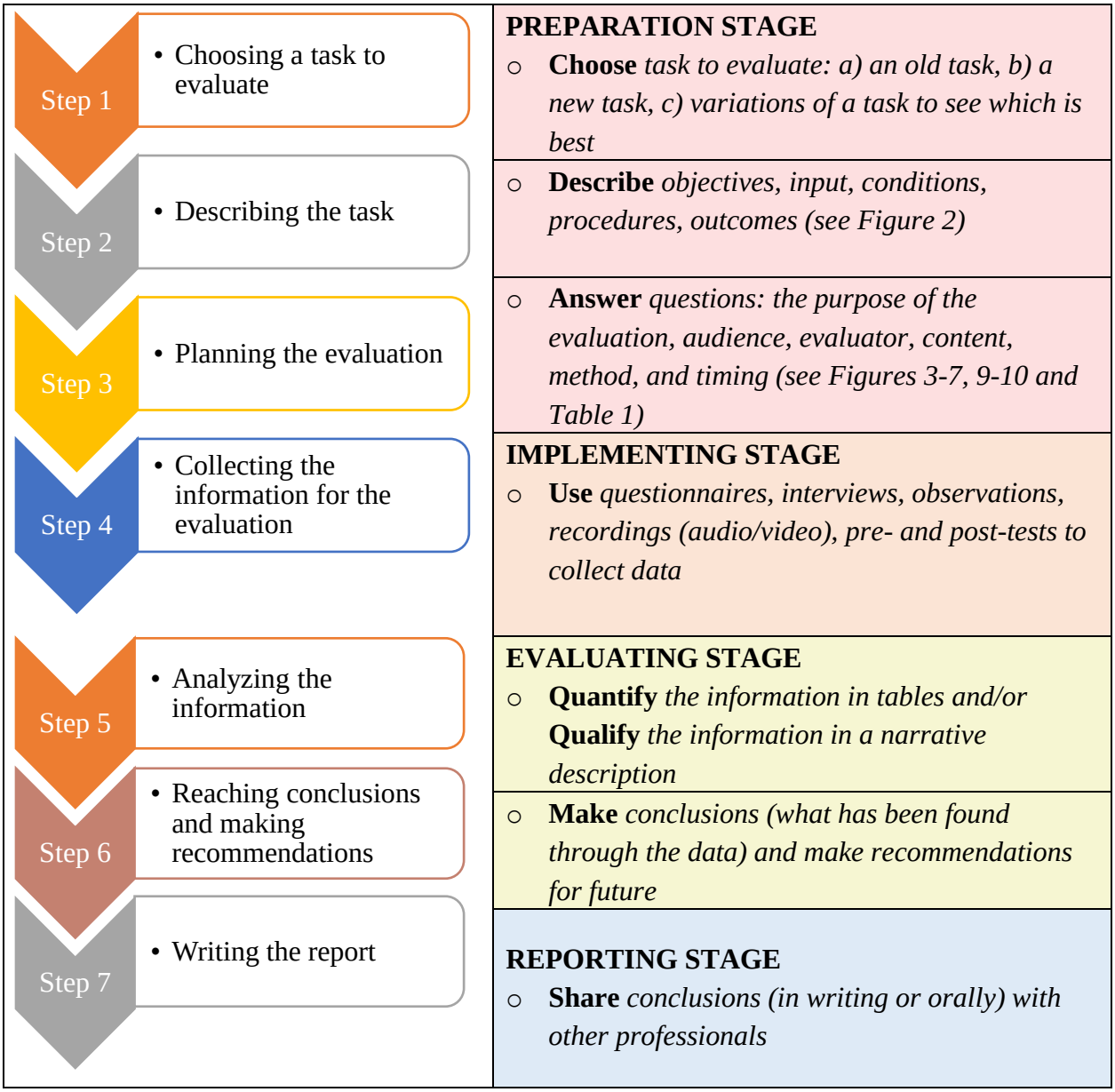


Figure 1. Seven steps to evaluating a task (Ellis, 1997, p. 38).

These steps make the process of evaluating a task easier to comprehend by breaking it down into distinct stages and steps to be undertaken in a certain order.

The Preparation Stage

The preparation stage of the task evaluation is made up of three steps as shown in Figure 1. Step 1 is choosing the task to evaluate. Step 2 is describing the task. Step 3 is answering evaluation questions about the task. Each step is described below.

Step 1: Choosing the Task

Each researcher starts the evaluation by selecting one task that they most want to research. The task could be selected because the teacher wants to check the effectiveness of a task he or she has been using, or check the effectiveness of a new task that the teacher wants to try out, or to check several variations of a task to see which variation is best.

Step 2: Describing the Task

Ellis (1997, p. 38) next explains how to describe a task. He states that a teacher/researcher creates a task description by describing *objectives, input, conditions, procedures,* and *outcomes* for the task (see Figure 2). The teacher/researcher first lists the *objectives,* the educational goals of the task; the *input*, what materials, information, or instruction the students receive to prepare for the task; the *conditions*, the layout and relevant facilities of the classroom, the grouping of the students in whole class, pairs, or small groups, the materials students will have or use for the task, whether the task is scripted or non-scripted, memorized or not, the time for preparation and doing the activity, and what is done after the task, etc.; the *procedure* for completing the task, described in a step-by-step manner; and the *outcomes* of the task in terms of both product and process; that is, what product should be produced as a result of the task, and what skills or language habits might be learned or what personal characteristics, e.g., the ability to work collaboratively, could be enhanced by going through the process of the task.

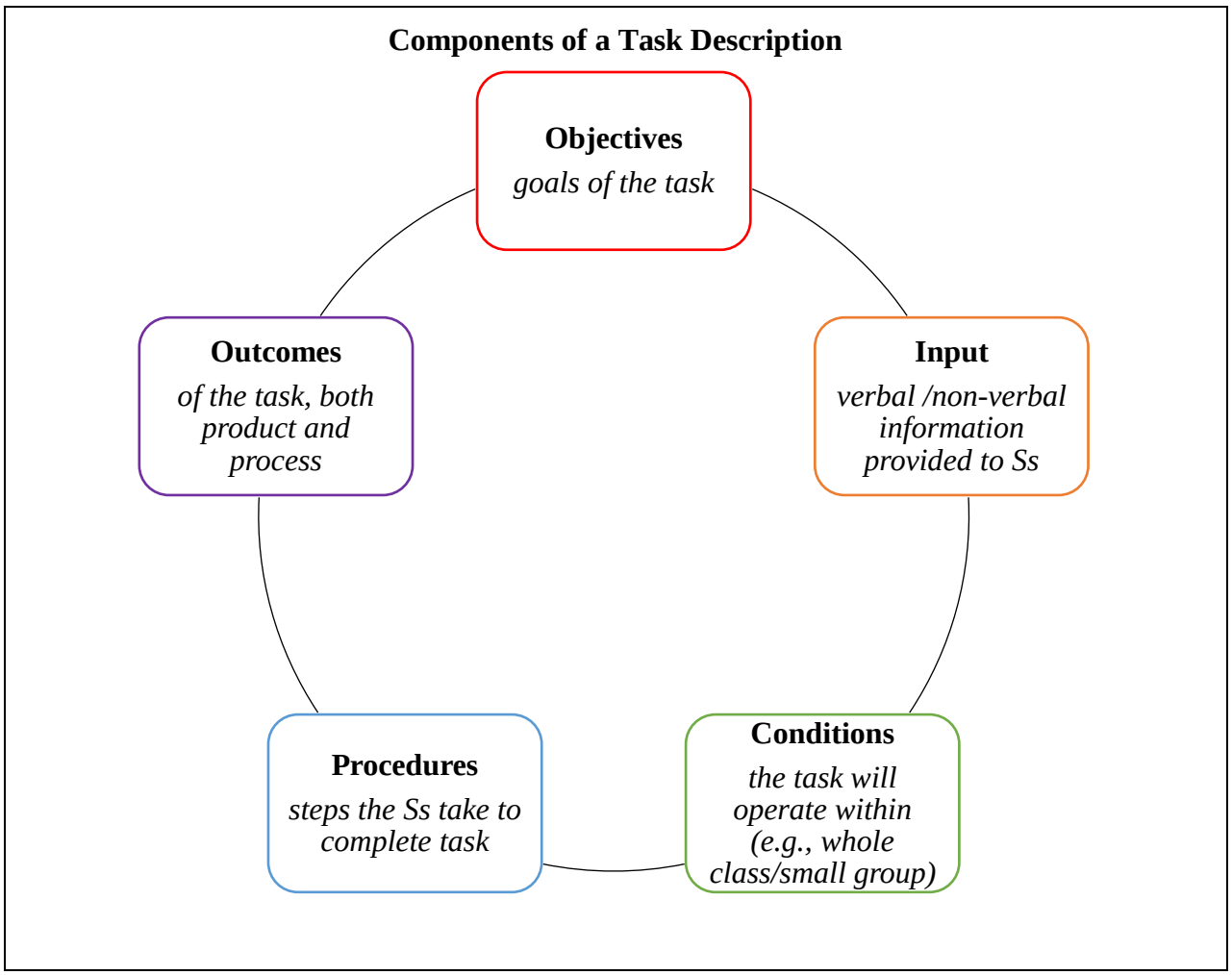


*Figure 2***.** Components of a description (from Ellis, 1997, p. 38).

Step 3: Answering Questions to Make Choices

The last step (Step 3) in the Preparation Stage (Steps 1-3) is to answer questions and make choices about the task evaluation (see Figure 3).

1. Purpose?		**2. Audience?**		**3. Evaluator?**		**4. Content?**		
Task met objectives?	Task can be improved?	Conducted for self?	Conducted for others?	Self?	Other?	Student attitudes?	Outcomes?	Learning?

5. Method?				**6. Timing?**			
Documentation?	Tests?	Observation?	Self-report?	Before task?	During task?	After task?	
						Immediately after?	*After a period of time?*

Figure 3. Questions and choices to make (Ellis, 1997, p. 39).

The questions are related to the purpose of the task, the audience for the evaluation, who the evaluator or evaluators will be, the content, the method used to evaluate, and the timing of the evaluation. In the next step the teacher/researcher will answer questions on the purpose, audience, content, method, and timing of the task evaluation, as outlined in Figure 3 and described in greater detail in Figures 4-10.

What is the purpose of the task? (See Figure 4.) The teacher/researcher can choose to use an objectives model which examines whether the objectives of the task were met, or a developmental model which looks at how the task can be developed and improved. It is possible to do a dual-purpose evaluation, objectives and developmental, with the developmental part described in the conclusion of the evaluation write-up.

1. Purpose?	
What is the purpose of the evaluation? What does the teacher/researcher hope to learn by doing the task evaluation?	
Task met objectives?	Task can be improved?
Were the objectives of the task met? (an objectives model evaluation)	In which ways can the task be improved? (a developmental model evaluation)

Figure 4. Explanation of purpose description of task (based on Ellis, 1997, p. 39-40).

Who is the audience for this evaluation? That is, who will be the beneficiary of the results of the evaluation? (See Figure 5.) The task evaluation could be done for the purpose of self-improvement as a teacher, or it could be written to be published so that other teachers could also benefit.

2. Audience?	
Who is going to learn about the task through the evaluation?	
Conducted for self?	Conducted for others?
Is the teacher/researcher the only person to learn about the results of the task evaluation?	Will the task evaluation results be shared with other teachers?

Figure 5. Explanation of audience description of task (based on Ellis, 1997, pp. 39-40).

Who is the person evaluating the data? (See Figure 6.) If the teacher is the evaluator of the data, it is relatively easy in that no effort is needed to find, train, and oversee outside evaluators, but having one or more outside evaluators will reduce the possibility of the evaluator being influenced by the teacher's opinions of individual students or by pre-conceived notions regarding the task. Having multiple external evaluators also increases the trustworthiness of the evaluation.

3. Evaluator?	
Who will be the evaluator of the task evaluation data?	
Self?	Other?
Will the teacher/researcher be the evaluator? (a common, easy way to conduct the evaluation)	Will an outside person be the evaluator? (more time-consuming as the evaluator would need to be trained, but results are often more trustworthy)

Figure 6. Explanation of evaluator description of task (based on Ellis, 1997, pp. 39-40).

What is the content of the evaluation? What part of the task will be focused on in the evaluation? (See Figure 7.) All three aspects of the task – student attitudes to the task, task outcomes, and learning outcomes – are important for the teacher to investigate. The teacher/researcher can choose any one of the task evaluation types, but by incorporating more than one of the three types, more data regarding the task is provided, and a more detailed evaluation of the task is possible.

4. Content?		
About the task, what is being explored?		
Student attitudes?	Outcomes?	Learning?
How interesting, useful, effective, easy, etc. did the students find the task? **(student-based evaluation)** These are the most common evaluation as it is the easiest. Usually uses questionnaires or interviews.	To what extent did the outcome of the task match the predicted outcome? **(response-based evaluation)** Usually uses live observation (with note-taking), audio recordings (transcripts), or video recordings (transcripts). Are time-consuming but result in useful data about the task.	Did the students learn anything, including the targeted objectives from the task? **(learning-based evaluation)** Often this requires pre- and post-task tests. Is the most difficult of the evaluations, but the resulting data is often more trustworthy.

Figure 7. Explanation of content description of task (based on Ellis, 1997, pp. 39-40).

More details about the three types of task evaluation can be found in Table 1 and Figure 10 below:

Table 1. Detailed Description of Three Contents of a Task (Ellis, 1997, pp. 39-40)

Evaluation Type	***Description***	***Investigation Method***	***Value***	**Problem**
Student-based evaluations	Students' attitudes to the task are examined. Did they find the task enjoyable, interesting, and/or useful?	Short questionnaires, interviews, focus group sessions with students. (One variation would be to do the same short questionnaires, interviews, focus group sessions with teachers who use the same task in their classrooms).	Gives important information on what students think about the task. Are the easiest kind to carry out.	Does not provide information on whether the objectives of the task were met, or whether any meaningful learning took place.
Response-based evaluations	Teacher examines the actual outcomes (both the products and processes of the task) to see whether they match the predicted outcomes.	Observations, live or recorded (audio or video), are examined to see if students are doing what the teacher/researcher intended them to do with the task.	Provide valuable information regarding whether the task is achieving what it is intended to achieve.	Time-consuming, demanding work, does not indicate whether meaningful learning took place.
Learning-based evaluations	The teacher/researcher attempts to determine whether the task has resulted in any new learning.	Pre- and post-task evaluations through tests, comparison of task before and after instruction.	Measures the learning that has resulted from performing a task.	*Most difficult to do. *May be difficult to measure the learning that has resulted from performing a single task.

A graphic illustration of the table can be seen in Figure 10 below.

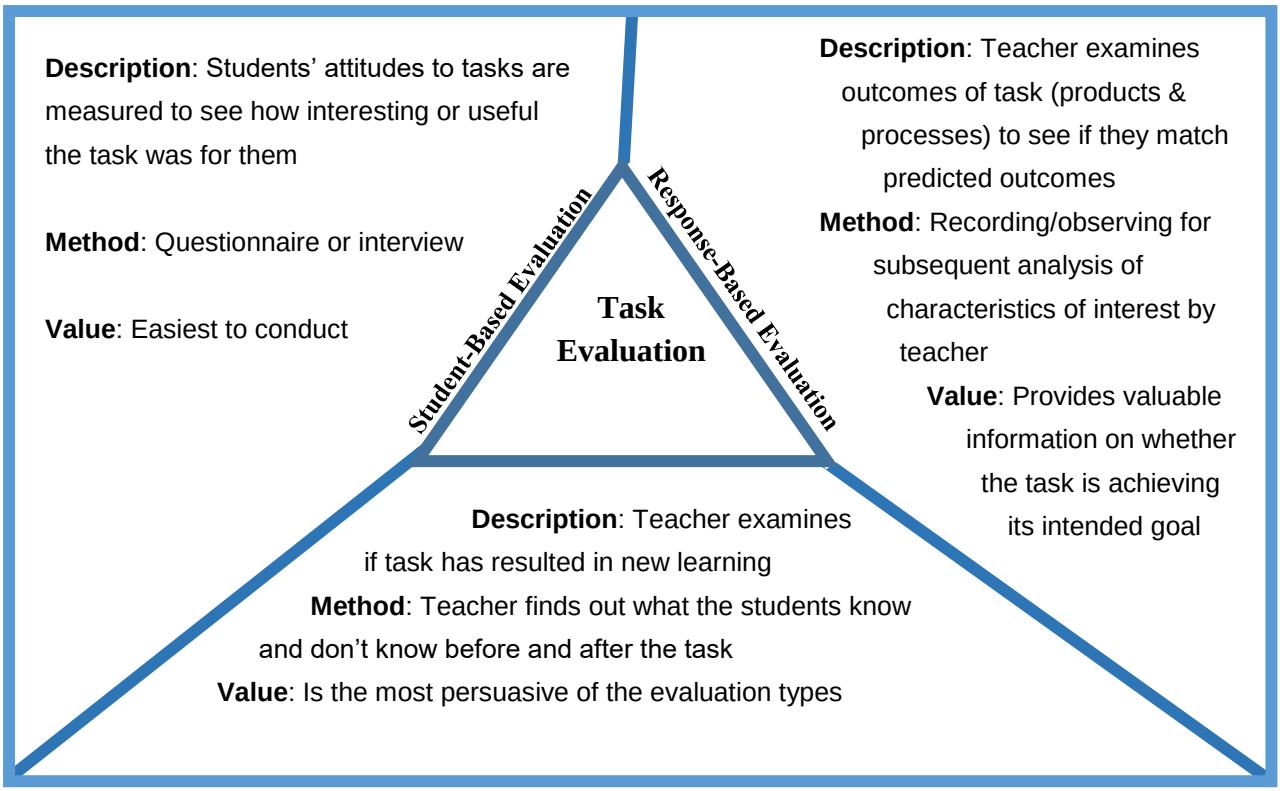


Figure 8. Types of task evaluation (modified from Ellis, 1997, pp. 39-40).

Ellis (1997, pp. 39-40) notes that student-based evaluations, typically using questionnaires to discover student attitudes toward the ease, usefulness, and entertainment value of a task or project, are the most usual form of research done in classrooms. This is because they are relatively easy to do, and more complex types of data are more difficult to obtain and analyze. However, student-based evaluations often do not supply data that would lead to sufficiently persuasive conclusions. It should be noted that a task could be evaluated through any combination of these three evaluation types, including the possibility of using all three types, that would supply enough data that could result in conclusions that might persuade administrators and other teachers. The more fully the task is evaluated, usually the more trustworthy the conclusions are, if the evaluation is done properly.

The kinds of evaluation instruments include documentation (e.g., student-written compositions), tests (pre-, post-, and formative), observation, and self-report (written or spoken). (See Figure 9.) The kind of evaluation instruments would depend on the content decided upon in Figure 7 above.

5. Method?			
What kind of evaluation instruments were used in the task evaluation?			
Documentation?	Tests?	Observation?	Self-report?
Were documents used (e.g., compositions from a writing task)? Used mostly for response-based or learning-based evaluations	Were tests used (written or oral)? Used mostly for response-based or learning-based evaluations	Were students observed? (live or audio- video-recorded) Used mostly for response-based evaluations	Did students report by themselves (self-report interview or questionnaire) Used mostly for student-based evaluations

Figure 9. Explanation of method description of task (based on Ellis, 1997, pp. 39-40).

What is the timing of the evaluation? That is, when will the data be collected? Ellis (1997) explains that the data to evaluate the performance can be collected prior to the task being performed (e.g., a survey on attitudes towards particular types of task or a pre-test), while it is being performed (e.g., observation, audio/video recording, performance as test), and after it is completed (e.g., surveys, written reactions to the task, interviews, focus group discussions, or a post-test), or any combination of these data collection types. If the data is collected after the task, the researcher has to decide whether the data will be collected immediately after the task or after a period of time has elapsed since the task. Students could also write a journal that would cover all three time periods: before the task, during the task, and after the task. (See Figure 10.)

<table>
<tr><th colspan="4">6. Timing?</th></tr>
<tr><td colspan="4">When will the task be evaluated?</td></tr>
<tr><td>Before task?</td><td>During task?</td><td colspan="2">After task?</td></tr>
<tr><td>Will Ss be given a pre-task interview or test?</td><td>Will Ss be observed during the task?</td><td colspan="2">When the task is ended, will the Ss be interviewed, given a questionnaire to fill out, or tested?</td></tr>
<tr><td colspan="2"></td><td>Immediately after?</td><td>After a period of time?</td></tr>
<tr><td colspan="2"></td><td>Will the Ss be evaluated immediately after the task?</td><td>Will the Ss be evaluated after a lapse of time after the task?</td></tr>
</table>

Figure 10. Explanation of timing description of task (based on Ellis, 1997, pp. 39-40).

The Implementation Stage

After the Preparation Stage, composed of steps 1 to 3, is completed, the Implementation Stage begins. Step 4 is to collect the data using the instruments that were decided upon (Figure 8) according to the timing that was decided upon (Figure 9). The teacher/researcher will collect the data (through questionnaires, interviews, observations, focus groups, audio or video recordings, or pre- and post-tests). The teacher/researcher is expected to abide by research ethics, and so will need to distribute bilingual consent forms to be completed by the teacher, student participants, and perhaps even by the institution.

The Evaluation Stage

The data that was collected is analyzed and arranged in a way that makes it easy to understand (Step 5). Quantify the information in tables and do a statistical analysis and/or qualify the information in a narrative description. Then make conclusions based on what was found in the data and make recommendations for the future (Step 6). If the teacher/researcher has questions or feels incapable of doing the statistical analysis, then the advice of fellow teachers should be sought.

The Reporting Stage

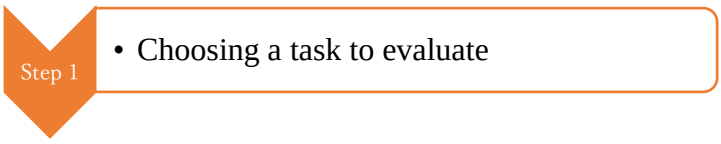
Step 7 makes up the Reporting Stage. The teacher/researcher shares the evaluation with other teachers orally in conferences or in writing in journals, unless the teacher decides to do the evaluation only for his or her self. However, it is advisable for teacher/researchers to write articles to be published or prepare a conference presentation to add to the knowledge base of the profession and provide the details for other teacher/researchers to replicate the evaluation in order to verify the results and conclusions.

Proposal: Applying the Task Evaluation Model to Performance-Assisted Learning

This proposal to address the lack of research studies supporting Performance-Assisted Learning (PAL) is to encourage PAL practitioners to become teacher-researchers and research what they are doing in their classrooms, and, most importantly, to write up their research. As mentioned earlier, Ellis (1997, p. 37) concludes, "A series of micro-evaluations can provide the basis for a subsequent macro-evaluation." That is the purpose of this proposal – to encourage the conducting of a large number of micro-evaluations of PAL that in aggregate will form a macro-evaluation of performance in learning. The way a PAL task micro-evaluation could be done is described below (cf., Figures 1-10, and Table 1).

Example Preparation Stage

As mentioned above, the Preparation Stage of a PAL task would be comprised of three steps. Step 1 is choosing a task to evaluate. Step 2 is describing the task. Step 3 is planning the evaluation. All three steps of the Preparation Stage are described below using the PAL task evaluation by Yoko, a PAL practitioner and teacher/researcher.



The task could be a speech, drama, debate, oral interpretation, or any other performance activity done in class. In this case, Yoko, the teacher/researcher, selects a variation of a speech activity she often has her students do – an informative speech – to check whether the variation is an improvement over the previously evaluated informative speech task.

Step 2 • Describing the task

Yoko describes the informative speech task she wants to evaluate as follows in Figure 11:

Objective: To successfully deliver an informative speech.

Input: Yoko teaches the students details about the structure of the information speech (Introduction, Body, Conclusion, Question Period) and the seven key performance qualities of a successful informative speech she will be rating the students on: Smoothness (smooth but without memorization), Energy, Loudness, Look (eye contact), Smile, Posture, and Gestures (SELLS+PG).

Conditions: There are 12 second-year university English majors, 10 females and 2 males. They sit in two groups of 6, 5 females and 1 male in each group. Students in a group sit around tables arranged in the shape of a U with the open end of the U facing the front of the room so that they can work in pairs or groups of 3 or 6, yet all can easily see the front of the room where the performer stands. In the first half of the course when they did presentations, they had to perform without a practice session, but in the latter half of the course the teacher wants to see if there were significant changes in learning due to the practice sessions.

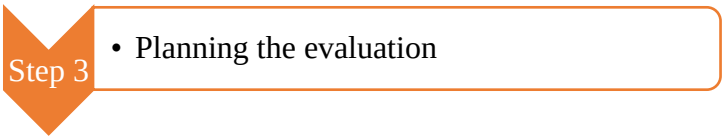
Procedure:

1. Each week, students create a poster in their B4-sized sketchbook on the topic of the unit as homework.
2. They practice a short speech on their poster as homework.
3. In class, after a general conversation warm-up (talking with a partner about what they did over the weekend), students individually move to an empty table, set their smartphones attached to a mini-tripod against the wall, set the smartphones to video, face their smartphones, and give their speech.
4. Students take out their earphones and watch and listen to their presentation while looking at a card with the organization and delivery items listed on it, noting what needs to be improved.

5. Students individually give their smartphone to an audience member to video record their presentation, which is done without notes and is not memorized, while the teacher also video records it.
6. At home, students watch their video and mark on an evaluation sheet their evaluations of the organization and 7 key characteristics of their presentation using a ten-point scale.
7. Students take a photo of their sketchbook page and evaluation and send them to the teacher.
8. The teacher watches the video and marks it on the same kind of evaluation sheet as students used.
9. Students write a final report on the information speech experience at the end of the semester.
10. Students complete an online survey about the activity.

Outcomes: Students should rate higher on organization and the seven key qualities when they are afforded a practice session than when they perform without a practice session.

Figure 11. Description of a sample informative speech task.



Yoko answers the questions regarding purpose of the evaluation, audience, evaluator, content, method, and timing of the task (see Figure 12):

Purpose of the Evaluation

Yoko wants to know if adding a self-recorded practice trial before doing the actual speech makes it easier to successfully perform the speech, changes the students' perception of their performance, and improves the actual quality of the speech. She wants to know if the task met the objectives and how the task can be improved.

Audience

Yoko wants to know herself but also wants to share the results of the evaluation with her colleagues at conferences and in papers.

Evaluator

Yoko would have liked to have had colleagues help her in evaluating the video recorded presentations, but because she did not ask earlier in the semester and the last part of the semester is extremely busy, she decides to be the evaluator. However, in the future she might ask other teachers to evaluate the recordings.

Content

Yoko decides to do a complete evaluation and decides to evaluate student attitudes, outcomes, and whether learning took place.

Method

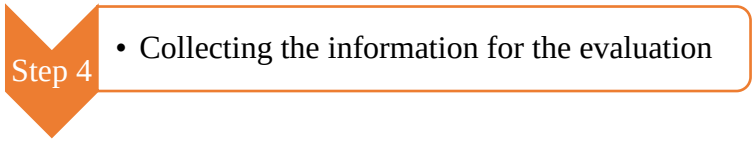
Yoko will use documentation (the final report), tests (the first pre-practice recording at the beginning of the semester and the last post-practice recording from each student), observation of partial transcripts from the recordings, and self-report (the students' weekly self-evaluations and an online survey of what students thought about the information speech activity.

Timing

Yoko will evaluate the task immediately after each speech, and a week after the final speech.

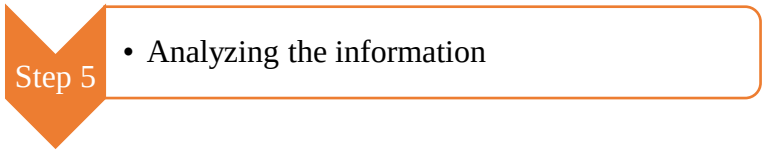
Figure 12. Example of planning an evaluation.

Example Implementation Stage



Yoko creates the online survey and collects the data, inputting it into spreadsheets for analysis.

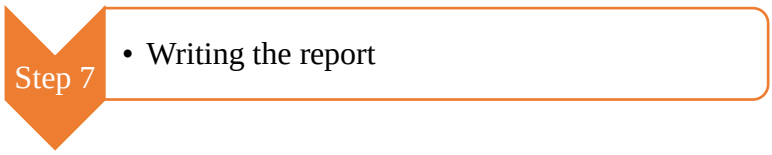
Example Evaluation Stage



Yoko uses simple statistical analysis (mean, median, high and low scores, and standard deviation) to analyze the quantitative data and identifies key concepts in the qualitative data.

Yoko looks at the data and determines that the practice had a strong effect on student attitudes toward the activity, helped improve the quality of the presentations markedly, and helped students to learn the key characteristics of performing a good information speech as well as key phrases in the conclusion and question and answer sections of the speeches. She recommends that video-recorded in-class practice be a standard feature of speech presentations.

Example Reporting Stage



Yoko gives presentations at several conferences, including one international one, and writes up a report for the international conference.

Conclusion: From Many PAL Task Micro-evaluations to a Macro-evaluation of PAL

This paper suggests taking the Ellis (1997) model for textbook micro-evaluation and repurposing it for PAL task micro-evaluations as Yoko did. This will help start to create a macro-evaluation of PAL if many PAL teacher/researchers do the same with their own PAL tasks. In my efforts to do micro-evaluations around the country, I have run into insurmountable resistance from institutions about having an outsider conduct research on teachers and students. Instead of one person attempting the herculean task of conducting the macro-evaluation, it should be easier for teachers within the institution to conduct the micro-evaluation and contribute the results to a pool of researchers' efforts.

A joint project conducted by teachers throughout Japan should give support to teacher/researchers who want to evaluate tasks, should answer many questions about PAL, and should provide answers that might be used to support PAL activities. Such a project will provide support to teacher/researchers so that they will not have to conduct their research alone but will have a group of researchers willing to help give advice. Finally, it will improve both individual teaching practice as well as the teaching profession.

If a large number of these micro-evaluations on PAL tasks were conducted in a careful manner and then collected in one accessible place, e.g., an online journal or at a conference with a digital post-conference publication, then PAL teachers perhaps would not have to go through the situation that Barbee (2016) and Carpenter (2015) experienced of having to answer to skeptical administrators and colleagues without adequate research data.

References

Barbee, M. (2016). Drama in the L2 classroom: A defense and practicum. *Mask & Gavel, 5*, 5-22.

Carpenter, J. (2015, November). *Incorporating drama into oral presentations*. Presentation at JALT2015 Conference, Shizuoka.

Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT Journal Volume 51*(1), 36-42.

Head, P., Kluge, D., Lyons, C., Rees, G., White, J., & Bussinguer-Khavari, V. (2018). Performance-assisted learning: Enabling students to thrive and jive. In P. Clements, A.

Krause, & P. Bennett (Eds.), *Language teaching in a global age: Shaping the classroom, shaping the world*. Tokyo: JALT, pp. 233-241.

Kawakami, A. (2012). Drama on the global stage: Cultural competence and the role of drama techniques in the EFL classroom. *Mask & Gavel, 1*, 16-21.

Kluge, D., & Catanzariti, J. (2013). Creative presentation and performance in the language classroom. *ACADEMIA, 93*, 77-106.

Maley, A., & Duff, A. (1978). *Drama techniques in language learning.* In P. Ur (Series Ed.). *Cambridge English Language Learning Series*. Cambridge: CUP.

Maley, A., & Duff, A. (2011). *Drama techniques: A resource book of communication activities for language teachers*. (3rd Edition). In P. Ur (Series Ed.). *Cambridge Handbooks for Language Teachers*. Cambridge: CUP.

NOTE 1: Rod Ellis will elaborate on the topic of this paper in a plenary talk in the June 15-17 conference in Nagoya.

NOTE 2: If you are interested in participating in the research project described above, please contact the author at klugeresearch@gmail.com.

NOTE 3: If you are interested in the research project described above, consider attending the June 15-17 conference in Nagoya where like-minded people will discuss the PAL task micro-evaluation project. See <https://sites.google.com/view/sddsigconferences/home>.

David Kluge (Nanzan University) has been teaching English for over 35 years. His research interests include oral interpretation, speech, drama, debate, composition, and materials development. He has co-authored with Matthew Taylor three books on composition (National Geographic Learning) and one book on oral communication. He was the founder and first coordinator of the Speech, Drama, & Debate SIG of JALT.