

Beyond the Rubric: Classroom Assessment Tools and Assessment Practice

November 2021 – Volume 25, Number 3

MaryLou Vercellotti

Ball State University, USA

<mlvercellott@bsu.edu>

Abstract

Analytic rubrics are promoted as important tools to assess learner performance and to improve learning outcomes. Rubrics, however, are not appropriate for every classroom assessment, particularly given the time and effort required to construct well-designed rubrics. In classroom assessment, instructors must balance the beneficial consequences of the assessment with the practicality to construct and implement the assessment in a timely manner. Informed by theory and empirical studies, this article reviews various assessment tools and practices that benefit learning. After a critique of analytic rubrics, it describes the strengths and weaknesses of three more practical tools for classroom assessment: checklists, scaled checklists, and detailed grading lists. Since the use of a more practical assessment tool might risk lowering the beneficial consequences of the assessment, the article then reviews three strategies which have been found to boost learning outcomes: discussing exemplars, providing effective feedback, and encouraging reflection. These assessment practices can be used in combination and with any assessment tool. The final section compares the tools for classroom assessment and summarizes how beneficial assessment practice can supplement the assessment tool, resulting in a balanced classroom assessment.

Keywords: *grading, scoring, assessment tools, practicality, washback, feedback*

Classroom assessment can be challenging, particularly because instructors often have to balance the competing ideals of having highly valid and beneficial assessments with reliable and practical assessment practice. Instructors often employ constructed response questions, where learners respond to assessment prompts, ranging from short responses (e.g., phrases, sentences) to long ones (e.g., paragraphs, essays, projects, portfolios) for classroom assessment (Hogan, 2013), and instructors typically use some type of assessment tool (e.g., answer keys or rubrics) to guide the assessment process. Each assessment tool should be well-matched for that assessment's purpose and task type (Brown, 2012). For instance, an analytic rubric is often a good fit for a longer performance-based task with a broad focus (e.g., general speaking ability). When the assessment task is more formative or narrowly focused, the assessment tool can be simpler. Regardless of the

assessment tool, assessment practice which engages the learner in the assessment process is likely to be more beneficial for learning.

The ultimate objective of classroom assessment is to benefit the learning outcomes (Green, 2020). Well-designed classroom assessments can benefit learning by influencing what happens both *before* and *after* the assessment. A classroom assessment should give clear information about the expectations of assessment, the *success criteria*, so learners can use that information to prepare for the assessment (Andrade, 2013). For example, when a learner knows a vocabulary quiz is scheduled, they often study those vocabulary words beforehand, boosting both learning and their performance. Likewise, sharing the success criteria for essay and project assessments can help learners improve their work before submission. Thus, good assessment practice requires that the assessment tool is given to the learners before the assessment so they will be more likely to submit work that meets the stated expectations (Ambrose *et al.*, 2010). Additionally, from a learning-oriented assessment perspective, teachers should help learners understand the success criteria (Turner & Purpura, 2016). After the assessment, the results give feedback to the learner about their strengths and weaknesses. In other words, the assessment process helps learners recognize the gap (or lack thereof) between their performance and the learning objectives which can be used to inform subsequent learning. In summary, assessment practice can increase the beneficial consequences of the assessment (Andrade, 2013), when the learner has detailed information about learning targets and receives detailed information about their strengths and weakness. An assessment that offers these benefits can often be less practical in that it takes more time and effort to create and implement. In classroom assessment, instructors must therefore find a balance between practicality and the benefits of a particular assessment within the specific teaching context.

The goal of this pedagogical research-informed article is to provide an overview of various assessment tools which can be useful in classroom assessment followed by an overview of assessment practices which can boost learning outcomes. It first reviews the main purpose and relevant shortcomings of rubrics, a commonly promoted assessment tool. It then reviews other assessment tools, with each subsection focusing on a different assessment tool. Each of these assessment tools is more practical than an analytic rubric but does not offer the detailed information of one, which means they give less support to the learners about what successful task-completion looks like. The third section suggests methods for building beneficial consequences into classroom assessment practice, regardless of the assessment tool. The final section offers a summary, including a comparison of the strengths and weakness of the reviewed assessment tools, and a recap of assessment practices which have been found to boost learning outcomes. By the end of this article, readers will know which assessment tools are well-matched for various assessment tasks and how to boost learning outcomes by incorporating specific assessment practices.

Shortcomings of Rubrics

Analytic rubrics are considered the gold standard in assessment (Suskie, 2009) and have become the go-to tool for all performance-based assessments (e.g., essays, presentations, and projects). They are used in all classroom teaching contexts, from elementary to college classrooms (Jeong, 2015). Analytic rubrics, however, are not the best assessment tool for every task (Wolf & Stevens, 2007). This assessment tool works well for longer language performances (such as presentations or essays) and when the purpose is to assess a more global skill, such as overall speaking or writing ability (Vercellotti & McCormick, 2021). For such tasks, the main purpose of rubrics is to help the instructor assess the quality of the work when there are no discrete correct answers. In addition to guiding the scoring, the use of an analytic rubric can benefit learning. First, analytic rubrics list specific categories connected to the stated learning objectives, making the expectations of the assessment clear. Second, analytic rubrics describe what successful performance looks like for each

success criteria, which helps illustrate any gaps between the expectations and the learner's performance for each in the rubric. (See Vercellotti & McCormick, 2021 for a review of analytic rubrics.) Analytic rubrics are expected to benefit learning because they combine these two features: explicit statement of the success criteria and descriptions of quality (Brookhart, 2018).

Nevertheless, the use of analytic rubrics in classroom assessment requires a balance between the beneficial consequences, which are high, and practicality of the assessment, which is low. Previously, assessment specialists created rubrics, but classroom instructors now often create rubrics for their own classrooms (Goldberg, 2014). Well-designed analytic rubrics are time-consuming to construct (Suskie, 2009; Wolf & Stevens, 2007). Poorly-designed rubrics can actually hinder learning by being too exact (which can reduce learner creativity) or too vague (Wolf & Stevens, 2007). In fact, instructors have difficulty outlining every attribute of a strong performance, and instructors admit that they rate learners on criteria that were not included in the rubric (Jeong, 2015). Certainly, if success criteria are not or cannot be clearly articulated in the rubric, the value of using analytic rubrics is weakened, in that they do not help learners understand how to be successful in the task. Further, although the use of an analytic rubric does likely improve the practicality during the grading process (Ambrose et al., 2010), a well-designed rubric can be impractical to create and use to grade assessments within a reasonable timeframe for each classroom assessment.

Analytic rubrics are dense with information in order to differentiate the performance along a continuum of quality. That density can become overwhelming, so much so that the learner may not benefit from detailed descriptions (Andrade & Du, 2005) because the learner may simply not read them or may not understand how to use them to improve. The analytic rubric can only boost learning outcomes when learners can make comparisons between their own work and the stated expectations, but learners do not always have this skill. Learners tend to be overconfident when they compare their work to the stated objectives (Guest & Riegler, 2021). Likewise, a long or overly detailed rubric can hinder the instructor's ability to focus on assessing fairly (Baryla, Shelley & Trinor 2012). Because the analytic rubric is overwhelming, the time and effort spent constructing a rubric for classroom assessment may be too high a price to pay if the rubric does not fulfill its promise of clarifying expectations and of supporting learning. Other assessment tools might offer a better balance between beneficial consequences and practicality for at least some assessments measuring productive skills.

Other Assessment Tools

Each assessment tool may have an application within a language classroom; the key is the match between the assessment goal and the assessment type. Instructors assess productive language skills with a range of tasks (e.g., short answers, essays, dialogues, monologues, projects) for a variety of pedagogical purposes, and so the instructor's classroom assessment toolbox should have a variety of assessment tools. In order to create effective assessments, instructors must know various options (Brown & Trace, 2016). This section reviews three practical assessment tools, checklists, scaled checklists and detailed grading lists, by describing the format of the tools, their potential uses, and their capacity to benefit learning.

Checklists

A checklist itemizes specific requirements of the task, and the instructor scores each item dichotomously (Brookhart, 2018), which means that checklists generally record task completion rather than differentiating the quality of the work. Each item on the checklist can correspond to a learning objective (e.g., use of standard writing conventions) or components of a learning objective (e.g., capitalization, punctuation). See Figure 1. A tally is a variation of a checklist, which records

the number of times (i.e., not just presence/absence) that the learner produces a particular task requirement (e.g., transition words used effectively). This assessment tool should not record errors, though, because a focus on errors may discourage production (Green, 2020), and therefore may hinder learning. Further, error counts ignore the severity of the error (Green, 2020). Instead, the tally checklist can be used to record the learner's demonstrated skill, rewarding successful attempts without penalizing unsuccessful ones.

Speaking Task: Request to Roommate	
<u>Sophia</u>	<u>Wei</u>
<input type="checkbox"/> Greeting	<input type="checkbox"/> Greeting
<input type="checkbox"/> Described reason for request	<input type="checkbox"/> Described reason for request
<input type="checkbox"/> Made request	<input type="checkbox"/> Made request
<input type="checkbox"/> Answered question	<input type="checkbox"/> Answered question
<input type="checkbox"/> Closing	<input type="checkbox"/> Closing

Figure 1. Example – Checklist

Since this assessment tool is a list of the success criteria for the task, checklists can benefit learning outcomes by supporting successful completion of the task (Brown & Harris, 2013). Ironically, the simplicity of checklists might discourage accurate self-evaluation. Checklists that require supporting evidence that proves adherence to the checklist requirements has been shown to be more effective than a checklist where learners only have to tick each box (Wood & Wood, 2018). Such *interactive checklists* are more likely to boost learning outcomes by requiring learners to engage more with the success criteria. Benefits to learning after the assessment tends to be quite low with this assessment tool in that the presence (checked) or absence (unchecked) of success criteria offers little feedback to learners.

Scaled Checklist

A scaled checklist (also called *simple rating scales*) lists the success criteria, and a way for the instructor to evaluate each criterion on a basic rating scale. By replacing a checklist's checkbox for a simple scale (Green, 2020), scaled checklists can gauge some measure of quality (not just presence/absence). The simple scale, however, does not describe the quality needed for the rating, as analytic rubrics do (Brookhart, 2018). Scaled checklists, therefore, are still less complex, which then allows the possibility of several success criteria to be assessed. Figure 2, for instance, lists ten success criteria with three levels of performance.

Speaking Task: Peer Discussion – Charity to Support									
Grading Key: 2 = needs improvement; 2.5 = approaching standard; 3 = meets standard									
	<u>Li</u>			<u>Tarik</u>			<u>José</u>		
Greeting	2	2.5	3	2	2.5	3	2	2.5	3
Stating Opinion	2	2.5	3	2	2.5	3	2	2.5	3
Asking Opinion	2	2.5	3	2	2.5	3	2	2.5	3
Polite Agreement	2	2.5	3	2	2.5	3	2	2.5	3
Partial Disagreement	2	2.5	3	2	2.5	3	2	2.5	3
Polite Disagreement	2	2.5	3	2	2.5	3	2	2.5	3
Reaching Decision	2	2.5	3	2	2.5	3	2	2.5	3
Extending other's Topic	2	2.5	3	2	2.5	3	2	2.5	3
Listener Support	2	2.5	3	2	2.5	3	2	2.5	3
Turn-taking Skills	2	2.5	3	2	2.5	3	2	2.5	3

Figure 2. Example – Scaled Checklist

Detailed Grading List

A detailed grading list describes each required element of the task performance with its point value. The list aligns with the learning objectives of the task (i.e., the success criteria) and any required elements. (See Figure 3.) A detailed grading list works well for smaller assessments with a few and for larger tasks with several success criteria. This tool can also be used with more objective assessments as well as to measure quality of performances in open-ended productive tasks, where differing point values represent differing levels of quality. This assessment tool is appropriate for both lower- and higher-stakes summative assessments. The detailed grading list in Figure 3 was created from the levels of performance descriptions in an analytic rubric (Vercellotti & McCormick, 2021), but the detailed grading list also includes a task requirement (time limit) which is less appropriate for inclusion in a rubric. Figure 4 shows a detailed grading list for a multimodal project-based assessment.

Without descriptions of performance describing how points are earned along a continuum of quality, this tool is easier to create for classroom assessment, and using points is easier than analytic rubrics to use when grading (Brookhart, 2018). Additionally, scoring can be more granular because this assessment tool allows varying point values and a scoring procedure with partial points. For instance, in Figure 3, the vocabulary category can be scored from 0-10 including ½ points in between the whole numbers. As a result, using a detailed grading list for an assessment gives the instructor flexibility during the scoring process to distinguish quality in learner work. Admittedly, the ability to assign partial points may result in inconsistent scoring, so the instructor should document how partial points are to be assigned during the grading process (or even before the scoring begins).

Formal Informational Speech		_____/50
<i>Required Elements:</i>	<i>Notes</i>	<i>Points Earned</i>
Vocabulary: Variety of words; word choice supports meaning and is appropriate for task		_____/10
Grammar Structures: Variety of complex structures and effective transition words; grammar choices support meaning		_____/10
Pronunciation: Both segmentals and suprasegmental patterns are understandable to all audiences		_____/10
Fluency: Pacing and pausing enhance message		_____/10
Content: Topic was explicitly stated; information was factual and relevant to the topic; specific details enhance message		_____/5
Length: Speech was within the time limits		_____/5
General Comments:		

Figure 3. Example – Detailed Grading List

There is a clear alignment between the success criteria and the detailed grading list for learners, and the detailed grading list states the expectations, without other levels of performance, which focuses learners’ attention to those requirements. With each learning objective (or part of each objective) assessed individually, this tool provides information to the learner about any gaps between the work and the stated expectations, but, as it does not describe levels of success for each criterion, there is less information about how to be successful. Further, given that a detailed grading list may only provide a score for each success criterion, the tool by itself gives little specific feedback to learners about how their work differs from the expectations. Beneficial consequences can be improved by having space on the assessment tool for feedback, either general comments (as in Figure 4) or both general comments and category specific comments (shown in Figure 3).

Presenting Research		_____/50
<i>Required Elements:</i>		<i>Points Earned</i>
Research topic and purpose clearly introduced; audiences' attention captured with a compelling "hook."		_____/5
Scholarly conversation about the research topic explained by briefly summarizing and synthesizing secondary research.		_____/10
Claim clearly stated about the topic and supported with relevant research.		_____/10
Presentation organized to support effective use of time.		_____/5
Multiple relevant modes of communication (i.e. visual, audio, text) used to skillfully convey ideas and support the presentation of research.		_____/10
Evaluation of peers' presentations included detailed feedback of multiple types (e.g., constructive criticism, reader response, questions). <i>Note: If fewer than 5 posts are made, full points cannot be earned.</i>		_____/10
Comments:		

Figure 4. Example – Detailed Grading List for Multi-modal Project

To summarize, checklists, scaled checklists, and detailed grading lists are tools with varying advantages and limitations. Given the range of pedagogical tasks used in classroom assessment, instructors can identify which assessment tool matches its purpose.

Strengthening Beneficial Consequences

As stated, classroom assessment can boost learning outcomes when the students understand the success criteria and receive sufficient feedback. Importantly, learners can be encouraged to take responsibility in meeting the learning objectives (Cheng & Fox, 2017) by engaging in the assessment practice. Learner engagement is an important part of assessment (Turner & Purpura, 2016). This section reviews specific assessment practices which have been empirically shown to strengthen the beneficial consequences: discussing exemplars, adding individual feedback comments, and encouraging reflection. These assessment practices can be used in combination and paired with any assessment tool. In each subsection, the assessment practice is explained and then justified with supporting research, followed by suggestions of the use of the practice in classroom assessment.

Discussing Exemplars

Providing exemplars or student examples (e.g., a student-created essays or portfolios), particularly with highlighted strengths and weaknesses, can help illustrate the expectations to the learner (Ambrose et al., 2010). Of course, instructors may be concerned that learners will copy from or mimic the exemplar, so the student examples should differ from the assessment in some way, such as topic (Hendry et al., 2012). Sowell (2019) suggested various ways instructors can guide learners to better understand the purpose of exemplars, such as by asking them to identify key features of the genre, and similarities and differences in the examples. As shown in Figure 5, the guide can ask learners to compare exemplars to identify the structure of a particular genre and to identify features of a more successful example. These instructor-led activities make it less likely learners will simply copy the example. Further, the instructor can decide whether the learners have access to the exemplars during all stages of the assignment or only limited access to the examples (Sowell, 2019), such as when introducing the assessment and reviewing the success criteria. Hendry *et al.* (2012) found significantly higher performance for learners who participated in a review of exemplars with a teacher-led discussion of the scoring rationale of multiple examples that ranged from weak to strong, but simply sharing the student examples did not improve performance. Exemplars are particularly needed in second-language classrooms because students may not have experience with the genre (Sowell, 2019) and the genre's rhetorical patterns may be quite different in their own culture (Kim, 2012). Given the possibility of learner differences in exposure to specific genres, the use of exemplars is a matter of equity, as well as pedagogical practice. In other words, the discussion of exemplars makes the expectations clearer to the learners, leveling the field, rather than expecting all learners to know the conventions of that genre.

Noticing the structure of this genre:
How are the examples similar?

- 1.
- 2.
- 3.

How are the examples different?

- 1.
- 2.
- 3.

Which example better meets the success criteria? Why?

Figure 5. Example — Questions to Guide Exemplar Discussion

Moreover, the instructor and the learners could use the student examples in order to identify the success criteria, which is more pedagogically powerful than simply providing the success criteria to the learners (Andrade, 2013). Although constructing an analytic rubric is challenging, learners may be more able to co-create a simpler assessment tool, such as a scaled checklist or detailed grading list. Since the benefit of exemplars seems to come from the instructor’s guidance connecting features of the exemplars to the stated expectations of the assessment, the assessment tool does not have to include all the information. A simpler assessment tool can outline the success criteria without overwhelming detail while the exemplars illustrate quality work.

Providing Effective, Individual Feedback Comments

Feedback (beyond a score or grade) is a common strategy to improve learning outcomes. Effective feedback is specific and task-referenced, rather than generic and learner-referenced (Green, 2020). For instance, “*Successful use of descriptive words!*” is detailed and directly related to the task while “*Good effort!*” is generic and referencing the learner (Green, 2020). Feedback should be understandable to students and actionable, meaning that the learners should have a better idea of what to do to be more successful. Table 1 offers some examples of task-focused positive and corrective feedback for the success criteria from Figure 1. The additional comments are particularly useful when the learner’s performance fails to meet stated expectations, earning only partial points. As Figure 3 shows, the assessment tool can have a dedicated space for such comments. Similarly, explanations or strategies for reaching the objectives has been found to result in larger learning gains (William, 2013).

Table 1. Example -Task-focused, Actionable Feedback Comments

Success Criteria	Positive Feedback	Corrective Feedback
Greeting	<i>Friendly greeting :)</i>	<i>Greeting was too formal for this task – speaking with a roommate.</i>
Described reason for request	<i>Reason for request was clear.</i>	<i>Reason was short; share more information when asking a favor.</i>
Made request	<i>Request included all necessary information.</i>	<i>Phrases, such as “I was wondering if...” makes the request more polite.</i>
Answered question	<i>Great use of ellipsis with direct answer. “Yes, I will __.”</i>	<i>Answer the question more directly.</i>
Closing	<i>Adding “thanks, again” in your closing was a great choice!</i>	<i>Closing was short which can seem impolite. What other phrases have we practiced?</i>

In addition, feedback must be timely (Ambrose *et al.*, 2010; Green, 2020), which does not always mean immediate. Actually, feedback that is too quick may hinder the learner’s reflection on their performance, but feedback about basic knowledge and understanding should be provided sooner (Ruiz-Primo & Li, 2013). Timely feedback is provided to the learner before the next assessment (on the same concepts) so that learners have the opportunity to improve their work (Green, 2020).

Accepting that providing individual feedback boosts learning but can be time-consuming, instructors can use strategies to make the process efficient. A class review of common issues is one potential strategy, and this review is likely more beneficial when the review is learner-led with the learners working together to resolve errors. This type of problem solving supports learning transfer (James, 2017). The use of technology can also save time in giving individual feedback. For instance, QuickMark in Turnitin allows instructors to save crafted feedback comments within an assessment and across assessments, which saves time, while also allowing personalized feedback (van der Hulst, et al., 2014). Van der Hulst et al. (2014) found that instructors estimated that using QuickMark reduced grading time by 10-30%. Even a simple word document can be used to write and save feedback comments to copy and paste as needed while using a learning management system (e.g., Canvas, Blackboard) in-document comment feature.

The use of technology can even improve the quality of the feedback. The instructor can connect feedback to specific phrases or sentences in the work, unlike summary comments, and pinpointing of specific parts of the learner’s work helps the learner implement the feedback (van der Hulst et al., 2014). Also, knowing that the comments may be used multiple times, instructors can take time to craft higher quality comments (van der Hulst et al., 2014). Given the pedagogical advantages of individual feedback, many instructors devote time to this practice, which can become impractical. If providing feedback is delayed, the beneficial consequences are diminished, so finding ways to simplify the assessment procedure is critical in classroom assessment. Technology and choice of assessment tool are two ways to boost practicality when providing individual feedback comments.

Encouraging Reflection on the Assessment

Simply giving time for reflection after the assessment supports learning (Ambrose et al., 2010; Suskie, 2009). Providing the opportunity for learners to reflect on their own learning is the main principle of assessment *as learning* (Cheng & Fox, 2017, p. 64). This reflection can be before

submission of the assessment or after receiving the results of the assessment.

Before receiving a grade or instructor feedback, learners could be asked to reflect on the quality of their work in relation to the success criteria. Lesley (2015) described a self-assessment checklist with the expectations which also asked learners to identify a strength and a weakness of their participation in a group discussion. An example which requires more specific reflection is shown in Figure 6. Along with the success criteria (in the middle column), there are columns for the learner to note how that objective has been met and a column for the learner to note what changes can be made to meet or exceed the stated objectives. Fluckiger (2010) has called similar guides *single point rubrics* [1] and stressed that their use is formative reflection. The purpose is similar to an interactive checklist in that the tool requires the learner to provide support for their evaluation, specifically connecting their work with the stated expectations. The success criteria were copied from the detailed grading list (Figure 3) which minimizes the time needed to create this reflection guide.

How I met the criteria	Meets Success Criteria	How I will revise to better meet or exceed the criteria
	Variety of words; word choice supports meaning and is appropriate for task.	
	Variety of complex structures and effective transition words; grammar choices support meaning.	
	Segmentals and suprasegmental patterns are understandable to all audiences.	
	Pacing and pausing enhance message.	
	Topic was explicitly stated; information was factual and relevant to the topic; specific details enhance message.	
	Speech was within the time limits.	

Figure 6. Example — Reflection Completed before Submission

After the assessment, engaging with any instructor’s feedback is crucial. Learner engagement with the feedback is the bridge between the instructor’s feedback and improved learning outcomes (Winstone et al., 2017). If learners do not read or understand how to act on the feedback, it will obviously not improve learning outcomes, regardless of the quality of it. Accordingly, instructors can create a post-assessment activity to encourage engagement with the feedback (Cheng & Fox, 2017; Winstone et al., 2017), which can be easy to add (Thompson, 2012). For instance, the reflection can ask learners to identify specific strengths and weaknesses, as shown in Figure 7. Cheng and Fox (2017) described a pair activity of “Thinking about doing better” where learners can list problems they see in their work and how to fix them. These types of guiding questions provide another opportunity to notice the gap between what they know and the target language. When given an opportunity, language learners tend to self-correct specific language forms (Vercellotti & McCormick, 2018), and self-correction activities have been found to be beneficial to subsequent language performances (e.g., McCormick & Vercellotti, 2013). Further, reflection on learning can facilitate noticing of patterns, which supports learning transfer (James, 2017), so learners can also be asked about language generalizations (e.g., patterns) and transfer (e.g., where can this learning be used in the real world) in the reflection (James, 2017).

What do I do well?

1. _____

2. _____

What should I review?

1. _____

2. _____

What questions do I have?

1. _____

2. _____

Where can I find the answers to my questions?

1. _____

2. _____

Figure 7. Example – Guide for Reflection after Assessment

Because learners often need support to process feedback, it is more effective if the reflection opportunities are offered during class meetings, encouraging learners to review and summarize the feedback and set learning goals (Ambrose et al., 2010; Ruiz-Primo & Li, 2013) rather than expecting learners to read, interpret, and implement the feedback independently. Without time for the learners to reflect, the instructor’s effort spent on feedback may be wasted which threatens practicality as well as the intended beneficial consequences. Since classroom meeting time is valuable, this assessment practice might be reserved for larger or more important assessments.

In summary, each of these assessment practices take time, but these practices have been empirically shown to boost learning, and learning is the ultimate purpose of assessment. The beneficial consequences of these practices can supplement or complement any assessment tool. Pairing a more practical assessment tool with the one of these assessment practices can provide a good balance between the benefits to learning and the practicality of the assessment procedure.

Assessment Tools and Practice Summary

Even though measurement tools are imperfect (Suskie, 2009), they have possible applications within the language classroom (Brown & Hudson, 1998). The key is the match between the assessment goal and the assessment tool. For instance, an analytic rubric would be an inefficient tool to use when assessing the proper capitalization in a writing sample, just as a tally checklist would be an ineffective tool to assess overall speaking ability. The purpose of the assessment (formative or summative) and the focus of the assessment (form-focused or broad skills) are critical when deciding which assessment tool to use.

As shown in Table 2, when comparing these assessment tools, a pattern emerges where assessment tools which tend to be more practical offer weaker beneficial consequences while assessment tools that tend to be less practical are more likely to benefit learning through more detailed expectations and feedback. Checklists (simple, tally, or scaled) offer a list of expectations without specifics of how to be successful. A detailed grading list provides more detailed success criteria than a scaled checklist but avoids the complexity of an analytic rubric. Given the concerns about an overwhelming analytic rubric, a detailed grading list may offer the best balance between practicality and beneficial consequences in how the success criteria is presented. The review showed that a

detailed grading list can be used in a variety of assessments, and this assessment tool offers flexible scoring with variable point values and a continuum of partial points able to be awarded.

Table 2. Summary of Tools for Classroom Performance Assessment by Least- to Most-focused on Quality

Tool	Strengths	Weaknesses
Checklist	<ul style="list-style-type: none"> • Simple list of expectations 	<ul style="list-style-type: none"> • Little focus on differentiating quality • Little chance for detailed feedback
Scaled Checklist	<ul style="list-style-type: none"> • Simple list of expectations • Distinguishes some quality of performance 	<ul style="list-style-type: none"> • Limited feedback about the quality of the work • Consistency is threatened with under-specified descriptions of quality
Detailed Grading List	<ul style="list-style-type: none"> • Detailed list of expectations • Can assess many requirements • Separate scores for each requirement • Scores distinguishes quality of performance with partial points 	<ul style="list-style-type: none"> • Does not describe the expected quality • Learners may not understand how their performance fails to meet expectations • Risk of inconsistency when awarding partial points
Analytic Rubric	<ul style="list-style-type: none"> • Distinguishes quality of performance • Separate scores for each category • Descriptions of each performance level 	<ul style="list-style-type: none"> • Creating detailed descriptions of each level of performance is time-consuming, challenging • Dense rubrics can be overwhelming for learners and instructors

A creative hybrid assessment tool may be designed when an assessment would benefit from the flexibility of a detailed grading list and clear descriptions of the expected quality of work at several levels of performance. For instance, an assessment may have a point system for completion of objective features and a one or two category analytic rubric which describes how to successfully meet the quality expectations.

Regardless of the assessment tool, the instructor can support learning outcomes by incorporating assessment practices which engage the learner (Brown & Harris, 2013). The underlying belief is that assessments can be learning tools in addition to measuring tools (Gezer-Templeton *et al.*, 2017). First, learning outcomes are improved when the connection between the assessment and the learning objectives is more transparent and explicit to the learners. Second, learners need support in understanding how to be successful, which can be accomplished with the assessment tool or with a combination of the assessment tool and assessment practice, such as using exemplars to illustrate how to be successful. Third, the assessment should provide learners with clear feedback of the gap between their work and the stated expectations. This goal can be accomplished with the combination of the assessment tool showing the results of the assessment and assessment practice giving feedback and the opportunity to reflect. The feedback comments and the reflection should be specific and focused on how the learners can meet the learning objectives. As shown in Table 3, any concerns about each of these assessment practices can be resolved by intentionally guiding the

learners to use the practice effectively, rather than expecting the learners to know how to successfully use exemplars, feedback, and reflections independently.

Table 3. Summary of Assessment Practices for the Classroom

Tool	Benefits	Concerns and Resolution
Exemplars	<ul style="list-style-type: none"> • Offer a model of how to be successful (especially important with cultural-dependent genres) 	<ul style="list-style-type: none"> • Learners may copy or mimic → Guide learners in the appropriate use of examples
Individual Feedback	<ul style="list-style-type: none"> • Specific, task-focused, timely, actionable feedback is very powerful method for improving learning outcomes 	<ul style="list-style-type: none"> • Time consuming → Use technology to improve efficiency • May be unread → Give time and guidance in implementing comments
Reflection	<ul style="list-style-type: none"> • Centers the learner as the benefactor of the assessment • Creates opportunity to process the instructor feedback 	<ul style="list-style-type: none"> • Time consuming → Modify existing documents • Reflection may be superficial → Encourage focus on content and implementation

Classroom assessment is a frequent and an on-going part of the teaching process, particularly in language learning. Every assessment should provide learners with explicit expectations, details of how to be successful, and feedback about strengths and weaknesses. The assessment tool does not have to provide all these beneficial consequences, however, because that attempt can result in an overwhelming and unwieldy classroom assessment tool. Additionally, learners should be engaged as partners in their own learning during the assessment process which facilitates robust learning. With several tools in their assessment toolkit, instructors can pair the right assessment tool with beneficial assessment practice to clarify the success criteria and provide sufficient feedback.

Note

[1] Since the format (no continuum of descriptors) and the purpose of a single point rubric (formative not grading) differs from an analytic rubric, this term seems like a possible source of confusion. [[back to article](#)]

About the Author

Mary Lou Vercellotti is an Associate Professor of Applied Linguistics at Ball State University. Dr. Vercellotti earned her PhD at the University of Pittsburgh. Her research interests include L2 language development as measured by language performance, prompt effects, assessment, English language pedagogy, professional development for teachers, and the Scholarship of Teaching and Learning.

Acknowledgement: The author wishes to thank Taylor Lutz, Dawn E. McCormick, TESL-EJ editor Kent Hill, and three anonymous reviewers for their helpful feedback on a previous version of this article.

To cite this article

Vercellotti, M. L. (2021). Beyond the rubric: Classroom assessment tools and assessment practice. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 25(3). <https://tesl-ej.org/pdf/ej99/a9.pdf>

References

- Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works: Seven research-based principles for smart teaching*. John Wiley & Sons, Inc.
- Andrade, H. L. (2013). Classroom assessment in the context of learning theory and research In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 17-34). SAGE Publications Ltd.
- Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation* 10(3). <http://PAREonline.net/getvn.asp?v=10&n=3>
- Baryla, E., Shelley, G., & Trinor, W. (2012). Transforming rubrics using factor analysis. *Practical Assessment, Research & Evaluation*, 17(4). <https://scholarworks.umass.edu/pare/vol17/iss1/4/>
- Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education*, 3(22). <https://doi.org/10.3389/educ.2018.00022>
- Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 367-393). SAGE Publications Ltd.
- Brown, J. D. (2012). Choosing the right type of assessment. In C. Coombe, P. Davidson, B. O'Sullivan & S. Stoyloff (Eds.), *The Cambridge guide to second language assessment* (pp. 133-139). Cambridge University Press.
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Brown, J. D., & Trace, J. (2016). Fifteen ways to improve classroom assessment. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 490-505). Routledge.
- Cheng, L., & Fox, J. (2017). *Assessment in the language classroom: Teachers supporting student learning*. Palgrave.
- Fluckiger, J. (2010). Single point rubric: A tool for responsible student self-assessment. *The Delta Kappa Gamma Bulletin*, 76(4), 18.
- Gezer-Templeton, P. G., Mayhew, E. J., Korte, D. S., & Schmidt, S. J. (2017). Use of exam wrappers to enhance students' metacognitive skills in a large introductory food science and human nutrition course. *Journal of Food Science Education*, 16, 28-36. <https://doi.org/10.1111/1541-4329.12103>
- Goldberg, G. L. (2014). Revising an engineering design rubric: A case study illustrating principles and practices to ensure technical quality of rubrics. *Practical Assessment, Research & Evaluation*, 19(8). <https://scholarworks.umass.edu/pare/vol19/iss1/8/>
- Green, A. (2020). *Exploring language assessment and testing* (2nd ed.). Routledge.

- Guest, J., & Riegler, R. (2021). Knowing HE standards: How good are students at evaluating academic work. *Higher Education Research & Development*. <https://doi.org/10.1080/07294360.2020.1867516>
- Hendry, G. D., Armstrong, S., & Bromberger, N. (2012). Implementing standards-based assessment effectively: Incorporating discussion of exemplars into classroom teaching. *Assessment and Evaluation in Higher Education*, 37(2), 149-161. <https://doi.org/10.1080/02602938.2010.515014>
- Hogan, T. P. (2013). Constructed response approaches for classroom assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 275-292). SAGE Publications Ltd.
- James, M. A. (2017). A practical tool for evaluating the potential of ESOL textbooks to promote learning transfer. *TESOL Journal*, 8(2), 385-408. <https://doi.org/10.1002/tesj.279>
- Jeong, H. (2015). What is your teacher rubric? Extracting teachers' assessment constructs. *Practical Assessment, Research & Evaluation*, 20(6). <https://scholarworks.umass.edu/pare/vol20/iss1/6/>
- Kim, E. J. (2012). Providing a sounding board for second language writers. *TESOL Journal*, 3(1), 33-47. <https://doi.org/10.1002/tesj.2>
- Lesley, J. (2015). Evolving monitoring templates and formative feedback checklists used in self/peer-reflection. *New Directions in Teaching and Learning English Discussion*, 1(1), 18-36.
- McCormick, D. E., & Vercellotti, M. L. (2013). Examining the impact of self-correction notes on grammatical accuracy in speaking. *TESOL Quarterly*, 47(2), 410-420. <https://www.jstor.org/stable/43267798>
- Ruiz-Primo, M. A., & Li, M. (2013). Examining formative feedback in the classroom context: New research perspectives. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 215-232). SAGE Publications Ltd.
- Sowell, J. (2019). Using models in the second-language writing classroom. *English Teaching Forum*, 57(1), 2-13.
- Suskie, L. (2009). *Assessing student learning: A common sense guide*. Jossey-Bass.
- Thompson, D. R. (2012). Promoting metacognitive skills in intermediate Spanish: Report of a classroom research project. *Foreign Language Annals*, 45(3), 447-462. <https://doi.org/10.1111/j.1944-9720.2012.01199.x>
- Turner, C. E., & Purpura, J. E. (2016). Learning-oriented assessment in second and foreign language classrooms. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 255-274). De Gruyter Mouton.
- Van der Hulst, J., van Boxel, P., & Meeder, S. (2014). Digitalizing feedback: Reducing teachers' time investment while maintaining feedback quality In A. Ørngreen, & K. Tweddell Levinsen (Eds.), *Proceedings of the 13th European Conference on e-Learning ECEL-2014* (pp. 243-250). Academic Conferences and Publishing International.
- Vercellotti, M. L., & McCormick, D. E. (2018). Self-correction profiles of L2 English learners: A longitudinal multiple case study. *TESL-EJ*, 22(3).

- Vercellotti, M. L., & McCormick, D. E. (2021). Constructing analytic rubrics for assessing open-ended tasks in the language classroom. *TESL-EJ*, 24(4). <https://www.tesl-ej.org/wordpress/issues/volume24/ej96/ej96a2/>
- Wiliam, D. (2013). Feedback and instructional correctives. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 197-214). SAGE Publications Ltd.
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist*, 52(1), 17-37. <https://doi.org/10.1080/00461520.2016.1207538>
- Wolf, K., & Stevens E. (2007). The role of rubrics in advancing and assessing student learning. *The Journal of Effective Teaching*, 7(1), 3-14.
- Wood, M. D., & Wood, J. M. (2018). Saving time, increasing learning: Using checklists to help students perform disciplinary writing conventions. *Journal on Excellence in College Teaching*, 29(2), 19-42.

Copyright rests with authors. Please cite TESL-EJ appropriately.