Rating Multilingual Learners' Written Language Consistently

We know from research and practice that all students need to develop writing skills. For multilingual learners the development of writing skills often lags behind their development in other language domains, particularly oral language proficiency. Nevertheless, writing skills are critical, as they are commonly associated with success in school (Gillespie et al., 2014; Graham et al., 2020; Puranik et al., 2014).

Educators can support multilingual learners and set them up for success in school by providing effective writing instruction and consistent writing evaluation. In this Focus Bulletin, we explore the process of evaluating student writing, and we examine how educators can consistently evaluate the written language of multilingual learners. We provide strategies for educators, working alone or with others in a community of practice, to support evaluations that generate focused and helpful feedback for students. We hope the topics covered here will help educators feel more

An open dialogue between experienced and newer educators is important to create a positive and supporting atmosphere in which new teachers can see the benefits of a collaborative approach to the evaluation of student writing.

confident in how they approach the process of evaluating student writing.

Why is it Important to Evaluate Writing Consistently?

Just as writing is a learned skill, so is the practice of evaluating writing. It involves understanding a wide range of factors, such as:

- Grade level knowledge that frames expectations for students' writing,
- The writing process students go through to demonstrate their writing skills,
- Socio-emotional factors between educators and students; for example, how a students has been participating in class recently.

Evaluating writing consistently is just one component of the process, but an important one. If writing is not evaluated consistently, the risks to students are real: to quality evaluation, to the feedback students receive, and ultimately, to the growth in students' writing skills. Let's consider how different approaches to evaluating writing may contribute to different outcomes for students and teachers.

Imagine two different schools, each one having several educators who all teach students how to improve their writing skills. In the first school, the teachers seldom discuss how they instruct their students, and they almost never compare with one another the writing performance of their students. They don't discuss among themselves what they value in their students' writing, nor do they have a common rubric that they use to evaluate that writing. Instead, the teachers all take their own approach.

In contrast, at the second school, the teachers all share what they value and look for in student writing. Together, they have created a rubric that defines these qualities and describes several score points. The teachers meet at the beginning of every school year to review writing performances from their students that they consider typical of each score point. They discuss these responses as a group, asking questions about points they may disagree with, and coming to group consensus on the scores the writing samples deserve. When new teachers come to the school, the more experienced teachers provide training in how to evaluate student writing.

It's clear in which context students are going to be better served. In the second school, students are more likely to receive evaluations of their writing that are consistent and clear. The teachers are better able to explain what a particular score means, in terms of skills a student has developed and demonstrated in their writing. Perhaps more importantly, the teachers understand which writing skills a student has not yet developed based on their writing performances and can tailor instruction to target the writing skills a student is ready to learn next. Evaluations of student writing depend much less on the subjective judgment of one person on a particular day, and more on the agreed-upon processes that have been developed and maintained over time by a group of educators.

Let's look at some steps you can put in place or build upon to help bring about a systematic approach to the evaluation of student writing that results in consistent and helpful outcomes. In

the context of assessment, we use the term "rating" when discussing the process of evaluating student writing.

Establishing Which Writing Skills are Valued

One important way to begin to evaluate, or rate, student writing is to establish which qualities in a text are valued at different points in the writing process. Early in the process, the focus may be on global aspects of the writing, such as how ideas are joined together both within sentences and across sentences. Later, the focus may shift to prioritizing vocabulary choices and control of sentence structures. Spelling may be particularly important in the production of the final product. It is important to specify these priorities, so that students and teachers are aware of which writing skills will be evaluated. Of course, these writing skills are likely to be the ones that are the focus of instruction in the classroom.

If we look through the lens of the WIDA English Language Development (ELD) Standards Framework, we can see the kinds of things that WIDA values in students' writing. For several years, the WIDA ELD Standards have articulated three dimensions of language: discourse dimension, sentence dimension, and word/phrase dimension. In the WIDA Writing Rubric and Scoring Scale, these three dimensions of language are the focus for how students' writing is evaluated. In the 2020 Edition of the WIDA ELD Standards, there is a particular emphasis on the discourse dimension, as can be seen in Figure 1.

While the WIDA ELD Standards Framework values students' ability to organize written language and to use cohesive devices, that does not mean every teacher and student needs to have the same priorities at all times. However, it is important that these priorities are clearly communicated

Figure 1. Grades 4-5 Expressive Proficiency Level Descriptors from the WIDA English Language Development Standards Framework, 2020 Edition

Criteria	End of Level 1	End of Level 2	End of Le
DISCOURSE Organization of language	Create coherent texts (spoken, written, multimodal) using...		
	short sentences linked by topic to convey an emerging sense of purpose (to inform, explain, argue, narrate)	sentences that convey intended purpose with emerging organization (topic sentence, supporting details)	short text tha intended purp using predicta organizational (signaled with paragraph or *first, and the*
DISCOURSE Cohesion of language	Connect ideas across a whole text through...		
	some frequently used cohesive devices (repetition, demonstratives)	some formulaic cohesive devices (pronoun referencing, etc.)	a growing n of cohesive (emerging articles to re same word, s antonyms)
DISCOURSE Density of language	Elaborate or condense ideas through...		
	a few types of elaboration (adding familiar adjectives to describe nouns: *maple syrup*)	some types of elaboration (adding newly learned or multiple adjectives to nouns *(thick, sweet, sticky maple syrup)*	a growing n types of el (adding art demonstrati nouns: *the da*

to students and to teachers who are involved in evaluating student writing. Ideally, these priorities will be documented and shared. In that way, there is clarity about how students' writing will be evaluated and what students need to do in order to receive credit.

It can be a worthwhile exercise for educators to collaborate on developing a rubric that specifies the writing skills to be evaluated, and if possible, describes different levels of performance for each skill. You may even be able to involve students in this activity. Building a rubric that is directly relevant to your local context, defines the writing skills that are prioritized, and describes these skills across several score points, is a very valuable activity, especially when it is done in collaboration with other stakeholders (teachers and students).

There is no shortage of guidance available online about developing a rubric. Here are two sites that we think are valuable: Yale University Poorvu Center for Teaching and Learning and the University of Minnesota Teaching with Writing website.

Establishing Benchmark Responses

Determining the writing skills that are valued within your local context is an important step towards rating students' writing consistently. A helpful activity is to create a set of benchmark student writing responses that are typical examples of each score point that you use in your context. These benchmark responses should exemplify the most commonly seen writing skills that you look for at each score point. These responses serve as the common standard that you and other teachers or raters look for to award that particular score.

The set of responses that you identify as your benchmarks can be created individually and will be relevant to the evaluations of your students' writing. If you are aiming to establish consistent writing evaluations for a larger group, perhaps within an entire school or district, you may want to consider group consensus on the writing benchmarks. A set of scored student responses that have been agreed upon by a group of educators can be a powerful tool for building a shared understanding of what each score point represents.

Once you have established a set of benchmark student writing responses, WIDA recommends writing short explanations about how each benchmark response exemplifies each score point. These explanations explicitly connect features of student writing to the wording used in the score point. The activity of creating these score point explanations, based on your benchmark responses, will help solidify the writing performance expectations for each score. Again, this activity may be done individually or collectively with a group of educators.

The Closer Look on page 5 shows some examples of score point explanations that WIDA has created. Please keep in mind these score explanations are intended for raters of student writing. They are not intended for students.

The explanations in the Closer Look provide justifications for the ratings given, and can be used to structure student feedback in a principled and consistent way. It is not necessarily meant to model that feedback. Instead, providing students with feedback that they can understand and practically use to improve their writing is very important.

CLOSER LOOK: Examples of Score Point Explanations

Protect students' privacy by redacting identifying information.

The explanations use descriptive language from the scoring scale along with quotes from the student's response. Together, the quotes and the language from the scale exemplify why the score point is the best fit.

Benchmarks Grade: 4–5
Task Name: Heating and Cooling

Response ID	Score Point	Explanation
Paper 4	1	This response is best described by the three descriptors at Score Point 1. It consists of minimal text [D1] that uses single phrases rather than complete sentences [S1]. Finally, it uses high-frequency English words (i.e., *help*) and words that are drawn primarily from the stimulus and prompt (i.e., *air conditioning*) [W1]. Thus, the holistic score for this response is 1.
Paper 11	2	This response is best described by the three descriptors at Score Point 2. It shows emerging organization (by using because to expand on the first idea) with heavy dependence on the stimulus and prompt [D2]. It uses only simple sentence structures [S2]. Finally, its vocabulary is drawn primarily from the prompt and stimulus (e.g., *ways of staying warm and cool, electric*) [W2]. Thus, the holistic score for this response is 2.
Paper 2	2+ (D2, S3, W2)	This response is best described by three descriptors from two adjacent score points. It resembles a list of simple sentences, relying heavily on the stimulus and prompt [D2]. It uses some complex sentence structures (e.g., *i see that in 1793 they yoused [used] a fireplace and hand-held fans.; i olso [also] see in 1910 they yoused [used] a radiator and a [an] electric fan.*) [S3]. Finally, it uses vocabulary that is primarily drawn from the stimulus and prompt (e.g., *fireplace, hand-held fans, radiator, electric fan, central heating system, air conditioning*) [W2]. Thus, the holistic score for this response is 2+.
Paper 17	4	This response is best described by the three descriptors at Score Point 4. It is organized and presents a clear progression of ideas (by including an introduction and conclusion, as well as transitional language, e.g., first reason, second reason, last but not least) [D4]. It contains complex sentences with grammatical errors that do not generally interfere with comprehensibility (e.g., *First reason of how they changes is because in the year 1793 they got warmed with a cheminni [chimney] and cool with a hand-held fans, and now they keep warm with hot air condition and cool with a fan and Last but not least in the year 1980 they had Central heating system to keep them warm, and to keep them cool they used air conditioning so they could have been cool are both complex*) [S4]. Finally, it uses a variety of vocabulary beyond the stimulus and prompt that generally conveys intended meaning (e.g., *compared, cheminni [chimney], reason[s], waste, money, paying, light traveling, cool fact*) [W4]. Thus, the holistic score for this response is 4.

If you work individually on evaluating student writing, the steps described above will most likely be sufficient to support consistent evaluation. Defining your scoring criteria (the writing skills that you value) and creating a set of annotated benchmark responses will take you a long way. You should also consider updating the benchmark responses regularly, perhaps annually or every other year, to keep these benchmark examples updated with changes in your students and their writing skills over time. For those who work with a group of educators on student writing evaluations, there are a number of additional steps that you can take to build rigor into your evaluation processes.

Going Beyond: Training and Certifying

Once you have established some of the processes and resources described, you may begin to think about formalizing some of the procedures even more. Especially if you are in a context where a team of educators works together on student writing evaluations, as new teachers join that team you may want to introduce them to the work you do to support consistent evaluations. You may wish to provide training on how you evaluate student writing. The table below highlights some specific features that may be included in such training to support consistent evaluation of student writing.

Step	Activity
Introduce the scoring criteria	Introduce and explain the writing skills that are valued in your context. Explain why these particular skills are valued and why other writing skills are less important.
Review the rubric	Provide the rubric you use in your context and some time to review it. Meet and address any questions that may come up following this review.
Review the benchmark responses	Provide the benchmark responses in a randomized order and ask for the responses to be sorted into rank order. Ask the new teacher(s) to talk through any responses that they were unsure how to order.
Complete a training set	Ask the new teacher(s) to score a set of responses, and compare their responses to the benchmark scores. Discuss those scores that differ from the true scores to help calibrate to the scoring standards.
Complete a certification quiz	Have the new teacher certify by rating either the training set, or if possible an additional set of responses. They should attain a passing score of at least 70% to meet the quality standards required to evaluate student writing.

Throughout the steps described here, an open dialogue between experienced and newer educators is important to create a positive and supporting atmosphere in which new teachers can see the benefits of a collaborative approach to the evaluation of student writing. Whenever possible, encourage newer teachers to explain their own thoughts on how they evaluate student writing. In many cases, new teachers may already be essentially on the same page as the team they are joining. They may just need a nudge here and there to bring them fully in line with how the established team evaluates student writing.

Sometimes, newer teachers may be anxious about activities that ask them to evaluate student writing for the first time, when their evaluations will be compared with benchmark scores. Provide encouragement by explaining that everyone provides feedback that differs from the group at first and that discussion of these responses where evaluations differ is very valuable, resulting in better calibration with established practices. Finally, even if you introduce a certification quiz with a passing score, it's not necessary for anyone to pass the first time! Taking the quiz a couple of times is normal and nothing to be concerned about.

And If You Need More …

The processes described in this Focus Bulletin will make a major contribution to the consistent evaluation of student writing. To sustain the quality you achieve, consider periodic monitoring of how student writing is evaluated. A monitoring system can ask teachers to meet periodically throughout the school year to collectively review and discuss evaluations of student writing. At these meetings, each teacher brings a small number of students' writing samples that have already been evaluated. These writing samples and the evaluations are shared with the group for discussion. It is preferable to redact student names from these writing samples so student privacy can be maintained. The aim of the meeting and discussion is to explore whether the evaluations are consistent across all teachers and that the scores and feedback provided to students reflect a shared understanding of the scoring criteria and the rubric. Effectively, these meetings are excellent professional development opportunities and help establish a professional community of practice in which student writing is evaluated consistently.
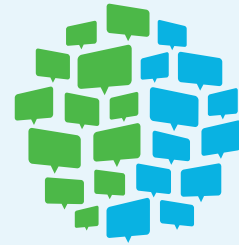
Additional Reading

There are many excellent books on how to evaluate student writing and the importance of evaluating students equitably. Some resources we would recommend are:

- *Assessment in the Second Language Writing Classroom* by Deborah Crusan

- *Assessing Writing* by Sara Cushing Weigle

- *Feedback in Second Language Writing* by Ken and Fiona Hyland

- *Assessing English language learners: Bridges to educational equity: Connecting academic language proficiency to student achievement* by Margo Gottlieb

References

Crusan, D. (2010). *Assessment in the second language writing classroom.* University of Michigan Press.

Gillespie, A., Graham, S., Kiuhara, S., & Hebert, M. High school teachers use of writing to support students' learning: a national survey. *Reading and Writing 27,* 1043–1072 (2014). https://doi.org/10.1007/s11145-013-9494-8

Gottlieb, M. (2016). *Assessing English language learners: Bridges to educational equity: Connecting academic language proficiency to student achievement.* Corwin Press.

Graham, S., Kiuhara, S. A., & MacKay, M. (2020). The Effects of Writing on Learning in Science, Social Studies, and Mathematics: A Meta-Analysis. *Review of Educational Research, 90*(2), 179-226. https://doi.org/10.3102/0034654320914744

Hyland, K., & Hyland, F. (Eds.). (2019). *Feedback in second language writing: Contexts and issues.* Cambridge University Press.

Puranik, C.S., Al Otaiba, S., Sidler, J.F., & Greulich, L. Exploring the amount and type of writing instruction during language arts instruction in kindergarten classrooms. *Reading and Writing 27,* 213–236 (2014). https://doi.org/10.1007/s11145-013-9441-8

Weigle, S. C. (2002). *Assessing writing.* Cambridge University Press.

WIDA Authors
Mark Chapman
Heather Elliott
David MacGregor

WIDA Focus Bulletin Staff
Rebecca Holmes
Rachele Stucker
Janet Trembley

Wisconsin Center for Education Research
University of Wisconsin–Madison
1025 West Johnson St.
Madison, WI 53706

Client Services Center
toll free: 866.276.7735
help@wida.us | wida.wisc.edu