

Amplify.

AWE Overview

Introduction

AWE is an automated writing evaluation tool that has been developed in conjunction with Amplify's regular writing activities, which ask students to use textual evidence to develop one focused idea or claim about the text and communicate that idea clearly and effectively to an audience. AWE is a tool designed to help teachers understand, track, and support student skill progress with these key foundational skills, which are a strong indicator of a student's analytic writing proficiency.

Overview: AWE, Core Writing Activities, and Foundational Skills

In the Amplify ELA lessons, students complete a writing activity 2–3 times weekly where they respond to a text-based prompt. The format of the prompts remains similar, so students can focus on their analysis of the text, and teachers can easily compare one piece of writing to another and track progress for three foundational skills that form the basis for strong analytic writing: Focus, Use of Evidence, and Conventions.

- Focus: To write exclusively about one moment or idea in order to fully develop it.
- Use of Evidence: Selecting, describing, and explaining quoted or paraphrased details from a text to develop and support an idea.
- Conventions: Use of grammar and sentence mechanics to control the clarity and power of sentences.

AWE provides auto scores for Focus and Conventions, allowing teachers to prioritize their assessment of students' progress with Use of Evidence, a skill that is relatively new for many middle grade students.

Note: Proficiency in each of these skills (Focus, Use of Evidence, and Conventions) requires that a student can maintain the skill over a period of writing. For this reason, the rubric score (and AWE score) for each skill is tied to word counts (writing productivity).

AWE Overview

Description: AWE Functionality

AWE Scores in Classwork

The AWE scores for Focus and Conventions appear in Classwork. The teacher will use a rubric to provide a score for Use of Evidence. An exclamation point and yellow warning triangle will appear if word count was too low and AWE was unable to generate a score for a piece of writing.

This screenshot shows the Classwork view of a writing activity that includes these auto scores:

ACTIVITY	SUBMISSIONS	LAST SUBMISSION	FEEDBACK						
7 WRITING Write Lesson 2	20/22	10/13/20 Thurs. 1:15pm	3 awaiting						
STUDENT	STATUS	LEVEL	FOCUS	CONV	EVID	CUSTOM SCORE	WORDS	COMMENTS	FEEDBACK
Anthony Bryk	Handed in 3/7/18 9:31am	▲▲▲	2	2	3	A-	128	3	
Mihaly Csikszentmihalyi	In Progress	●●●					99	0	
Carol Dweck	Handed in 10/13/20 11:37am	▲▲▲	1	1		B-	87	0	
Jamie Escalante	Resubmitted 10/12/20 7:04pm	▲▲▲	2	2		B+	71	1	
Fatima al-Filtri	Handed in 3/7/18 9:31am	●●●	4	4		C+	130	0	

AWE Scores When Viewing Student Work

When a teacher opens a piece of student writing the AWE scores given to that piece of writing will display. The small "AS" indicates that the score above is the AWE score for that skill.

Write a Body Paragraph
Dahl & Narrative Sub-Unit 4 Lesson 1

← Zimba Anthony Bryk Dweck →

Write about one of the red herrings you found. Explain why this detail seemed important and what it led you to predict.

As an informal fallacy, the red herring falls into a broad class of relevance fallacies. Unlike the straw man, which is premised on a distortion of the other party's position, the red herring is a seemingly plausible, though ultimately irrelevant, diversionary tactic. According to the Oxford English Dictionary, a red herring may be intentional, or unintentional; it does not necessarily mean a conscious intent to mislead.

The expression is mainly used to assert that an argument is not relevant to the issue being discussed. For example, "I think we should make the academic requirements stricter for students. I recommend you support this because we are in a budget crisis, and we do not want our salaries affected. The second sentence, though used to support the first sentence, does not address that topic.

Word Count: 132

SCORES COMMENTS

Custom Score
80

ASSESS SKILLS

Conventions
1 2 3 4
3 AS

Focus
1 2 3 4
4 AS

Use of Evidence
1 2 3 4

Send feedback

Overriding the AWE Auto Score

Teachers always have the option to replace the AWE auto score with their own assessment score. Simply click on the rubric score you identify for that skill and your score will replace the AWE score in Classwork. In the screenshot below the teacher selected a score of 3 for Focus, replacing the AWE score of 4. The "AS" tag persists, purely for the teacher's information. AWE does not score Use of Evidence. This skill is assessed by the teacher. The teacher can also provide a custom score to the student.

The screenshot shows a user interface with two tabs: "SCORES" and "COMMENTS". Under "SCORES", there is a "Custom Score" input field containing the number "80". Below this is a section titled "ASSESS SKILLS" with an information icon. It lists three skills: "Conventions", "Focus", and "Use of Evidence". Each skill has four circular buttons representing scores 1, 2, 3, and 4. For "Conventions", the score 2 is selected and has a blue "AS" tag. For "Focus", the score 3 is selected and has a blue "AS" tag. For "Use of Evidence", the score 1 is selected. At the bottom of the "SCORES" section is a purple "Send feedback" button.

Explanation: How AWE Works

AWE Score for Focus

The AWE focus-scorer assesses the degree to which a piece of student writing maintains Focus, with Focus defined as sustained attention to a single moment or idea. AWE will provide a Focus score between 1 and 4. Higher scores indicate that the student has written a relatively long response with many details relevant to a single moment or idea and few or no rough topic shifts (e.g., non sequiturs, tangents). Lower scores indicate a shorter response with a number of rough topic shifts.

Focus score and word count: Since the Focus rubric assigns scores relative to both word count and grade level, the scores are sensitive to word count and grade level.

AWE Focus Examples

SAMPLE 1: AWE Focus Score of 4

Grade 6 Response:

I agree with the idea that Mrs. Pratchett is a horror because she is filthy in the way she looks and she is also nasty with the boys who come into her shop. First of all, she has a greasy apron and food all over her blouse. She had, “toast-crumbs and tea stains and splotches of dried egg-yolk” on her apron. Can you imagine going to get a treat at the store and seeing the worker looking like that? Then, she uses dirty hands to get the candy for the boys. It even says she has “black fingernails” and she uses those nails to touch all the candy. She should wash her hands or use gloves. After all, candy is something you eat and you don’t want it to be all dirty. Finally, she is so mean she won’t even give the kids a bag unless they “spent a whole sixpence,” otherwise, she puts the candy in dirty newspaper. A sixpence was probably a lot of money and who wants newspaper all over their candy? She was clearly a horror.

AWE scores this sample as a 4 for Focus relative to grade 6:

- The response is over 120 words.
- All of the sentences refer to a central moment or idea.

SAMPLE 2: AWE Focus Score of 3

Grade 6 Response:

I believe Mrs. Pratchett is a horror because it says “Her apron was grey and greasy. Her blouse had bits of breakfast all over it,” She doesn’t care about the way she looks even though she is working at a store. Another reason I believe Mrs. Pratchett is a horror is because her hands were dirty. “They looked as though they had been putting lumps of coal on the fire all day long.” Her job is to touch the candy that people buy so her hands should be spotless. She is also mean. Instead of being friendly to her customers, she yells at everyone. Once I was at a store where the worker was being rude. I was trying to buy sneakers but she wouldn’t bring the right size so I left.

AWE scores this sample as a 3 for Focus relative to grade 6:

- The student has written over 120 words and is therefore eligible for a score of 4.
- The lack of focus displayed in the final part of the prompt lowers the student’s score to an overall 3.

SAMPLE 3: AWE Focus Score of 1

The AWE scorer recognizes that an essay consisting mostly of outright non sequiturs should be scored a 1. Consider the following 137-word text constructed using sentences from the Gettysburg Address and Wikipedia articles on *The Lord of the Rings* and Mickey Mouse.

Grade 6 Response:

Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battle-field of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. The One Ring slipped onto Frodo's finger inadvertently in the Prancing Pony's common room, turning Frodo invisible. This attracted the attention of Sauron's agents, who ransacked the hobbits' rooms in the night. The group, under Aragorn's guidance, quickly fled through the Midgewater Marshes and again escaped the Nazgûl. The Opry House (March 28, 1929) was the first time in which Mickey wore his white gloves.

AWE scores this sample as a 1 for Focus relative to grade 6:

- The student has written over 120 words and is therefore eligible for a score of 4.
- The large lack of focus and outright non sequiturs displayed in the response lower the student's score to an overall 1.

Note: The scoring algorithm assesses the extent to which the details in the writing remain relevant to a single moment or idea. For this reason, large sections of quoted text often may produce high focus scores. In these cases, teachers may want to instruct students on the importance of balancing quoted text with the student's analysis.

AWE Overview

AWE Score for Conventions

AWE will provide a Conventions score between 1 and 4. In general, those scores correspond to the writing behaviors described in the following table:

Grammar Score	Description
1	The high number of broken/ignored grammatical conventions would make the writing difficult for the reader to understand (basically too many errors per sentence).
2	Most sentences contain errors which impede the reader's ability to understand the writing OR a particular error (or set of errors) reappears with high frequency throughout the writing.
3	Most sentences are grammatical but there may still be a few errors. The writing is probably mostly clear (or if it is unclear, it is due to semantic or spelling errors).
4	Almost none of the sentences contain egregious errors. The writing is clear (or if it is unclear, it is because of semantic or spelling errors).

In scoring Conventions, emphasis is placed on the following kinds of errors (ordered roughly by decreasing emphasis¹):

- Run-ons
- Fragments
- Unusual or unparsable grammatical constructions
- Punctuation use
- Subject/verb number agreement
- Misspelled words/incorrect verb morphology
- Incorrect capitalization
- Excessive verb tense shift

Conventions score and word counts: Since the Conventions rubric assigns scores relative to both word count and grade level, the scores are sensitive to word count and grade level.

Note: This scorer was trained on student data, and it also uses language models trained on modern English extracted largely from news text, the Web, and Wikipedia. If given more complex analytic text (Shakespeare or a New Yorker essay—something that we might

¹ This emphasis order mostly comes from actually looking at the kinds of errors that we see in student writing.

AWE Overview

otherwise consider flawless English), it might score it as a 3 for reasons related to the uniquely low repetitions of word or construction types found in that kind of high-level text. As the algorithm matures and learns from “studying” more student writing, scoring will become more accurate.

Validity: How Well Does AWE Work?

These scores are trained on student behavior as seen across a wide spectrum of classrooms, which means that each student’s writing is being scored relative to a large number of students—not just those from his or her class. The scores are normalized for many students, not just the top performing ones.

Amplify currently expects essays to be scored “correctly” about 70% of the time. Amplify expects AWE scores to be off by one about 20–25% of the time. Amplify expects AWE to miscalculate by more than one in about 5–10% of cases. The accuracy of AWE will increase over time, as the algorithm works with increasing numbers of student writing pieces.