

Assessing Appropriate Assessments

Samantha Burleson

William Carey University
Mississippi Beta Chapter

Vol. 7(2), 2022

Title: Assessing Appropriate Assessments

DOI: 10.21081/ax0334

ISSN: 2381-800X

Keywords: assessment formats, high-stakes assessments, multiple-choice, critical response

This work is licensed under a Creative Commons Attribution 4.0 International License.

Author contact information is available from aletheia@alphachihonor.org.

Aletheia—The Alpha Chi Journal of Undergraduate Scholarship

- This publication is an online, peer-reviewed, interdisciplinary undergraduate journal, whose mission is to promote high quality research and scholarship among undergraduates by showcasing exemplary work.
- Submissions can be in any basic or applied field of study, including the physical and life sciences, the social sciences, the humanities, education, engineering, and the arts.
- Publication in *Aletheia* will recognize students who excel academically and foster mentor/mentee relationships between faculty and students.
- In keeping with the strong tradition of student involvement in all levels of Alpha Chi, the journal will also provide a forum for students to become actively involved in the writing, peer review, and publication process.
- More information can be found at www.alphachihonor.org/aletheia. Questions to the editors may be directed to aletheia@alphachihonor.org.

Alpha Chi National College Honor Society invites to membership juniors, seniors, and graduate students from all disciplines in the top ten percent of their classes. Active on nearly 300 campuses nationwide, chapters induct approximately 10,000 students annually. Since the Society's founding in 1922, Alpha Chi members have dedicated themselves to "making scholarship effective for good." Alpha Chi is a member in good standing of the Association of College Honor Societies, the only national accrediting body for collegiate honor societies. A college seeking a chapter must grant baccalaureate degrees, be regionally accredited, and be a not for profit institution.



Article Title: Assessing Appropriate Assessments

DOI: 10.21081/ax0334

ISSN: 2381-800X

This work is licensed under a Creative Commons Attribution 4.0 International License.

Assessing Appropriate Assessments

Samantha Burleson

William Carey University

Mississippi Beta Chapter

Abstract

This paper is an open discussion into the most beneficial assessment formats and how English Language Arts (ELA) teachers should decide which formats to use in their literature courses. ELA requires multiple assessment formats to measure student mastery. This paper defines different assessments, their purposes, and the criteria for each. Teachers must consider many variables when deciding. The results of this report may influence educational leaders to reevaluate their curriculum and question major tests.

Keywords: *assessment formats, high-stakes assessments, multiple-choice, critical response*

Assessments are the measure of success and failure in the education world. They range from pre-assessment, formative, summative, and high-stakes assessments, and their purpose is to measure a person's abilities. Education leaders rely on the results of high-stakes testing to change the support provided to schools, but in class, assessments are just as important to measuring mastery of standards and concepts for these high-stakes tests. This paper is an open discussion into the most beneficial assessment formats and how English Language Arts (ELA) teachers should decide which formats to use in their literature courses.

"Assessment" is a nonrestrictive term describing the different tests and measurements teachers use with their students (Popham 13). Popham states that an "educational assessment is a formal attempt to determine students' status concerning educational variables of interest" (12). Educational assessments allow teachers to ask students the same questions regarding any topic desired by the teacher. Tests are a tool used to measure the skill or knowledge taught (11). Educators use various assessments to measure student knowledge before a lesson, progression during a lesson, knowledge and new skills gained from a lesson, and school-wide measurement of progression during the school year.

Tailoring ELA classroom assessments is necessary to fit the class's interests as well as encompass the different style areas covered in this core course. Reading, writing, and grammar are three different focuses English teachers attack every day. Assessing these specific areas comes with challenges, such as learning abilities, accommodations, lack of prior knowledge, and disinterest. Deciding which assessment format fits into the course and purpose of the lesson is the biggest priority when teachers assess their students.

Pre-assessments measure knowledge and abilities before the lesson begins to inform teachers on what knowledge or skills students need to review or which ones to skip when the lesson begins. The pre-assessments serve as a baseline for the summative—or final—assessment to measure the percentage of growth from the class. This percentage assesses the students' progress along with the teacher's effectiveness during the lesson. Pre-assessments are a tool teachers use in the classroom to measure students' prior knowledge, while projected achievement level assessments provide data for school leaders to measure student growth from the beginning, the middle, and the end of school year. These project-

ed achievement level assessments act as a pre-assessment to state tests. Teachers do not always need written pre-assessments in their classrooms because they know their students and their abilities prior to planning instruction. However, at the beginning of the school year or when preparing for state and district tests, it is helpful to have students take pre-assessments for proper evaluation of prior knowledge and growth.

According to Hooley and Thorpe, "Molding student learning behaviors through formative assessment feedback has been shown to increase skill level, motivation, and metacognition" (1217). Formative assessments take place during a unit of instruction. Formatives provide the teacher with information on the students' progression through a unit, and from this information, the teacher modifies instruction to reach mastery. If the students are performing well on formatives, the unit ends early. If some students are struggling to grasp a concept, then the teacher changes his or her approach to teaching a concept. Quizzes, bell-work, writing activities, graphic organizers, class discussions, enter/exit tickets, in-class comprehension checks, and observations all serve as formative assessments in the classroom. Formatives are not just graded pieces of paper that assess gain of knowledge or attentions paid from a few lessons but are ongoing comprehension checks that a teacher does every day in the classroom. Conducting a small conversation or asking guided questions is one way a teacher gauges his or her students' grasp of a new concept. If a student cannot identify the theme of a story after the teacher explains how the elements work together to teach a lesson, then the teacher knows to change his or her approach. Reading comprehension or reading assigned for homework is part of formative assessments. Small quizzes asking for details in the story allow teachers to check if his or her students are retaining information from class or individual readings. Formative assessments help teachers and students communicate with each other about what is understood or misunderstood.

Summative assessments are final assessments measuring mastery of a standard, skill, or concept. This measure of mastery tells the teacher that it is acceptable to move on to the next standard since most skills and standards build on one another. Scores on summative assessments show the teacher who needs further explanation, which areas students struggle with, or if or that students have mastered a skill and are ready to move on. Summative assessments are also useful for teaching teams to

measure the growth of their students across grade levels. If there are multiple teachers for the same grade, then the teachers use their summative scores to measure their effectiveness of instruction within a unit. Teachers then use this information to modify their instruction and collaborate with each other. If one teacher's class performs well with a variety of activities or instructional methods, then the underperforming class's teacher can use that activity or method to reteach the unit or skill.

Classroom-wide assessments of are beneficial for high-stakes test preparation. All assessments, formal or informal, are teacher evaluations of student progress throughout the year leading up to state tests. Schools around the United States take different tests in separate districts, which allows them to measure both individual and state-wide growth. For example, Covington County School District in Seminary, Mississippi, takes the CASE21 tests to measure projected achievement level, growth, and proficiency. At the beginning of the school year, students take CASE21 tests in reading, math, social studies, and science to measure their current level of achievement and projected growth. Every nine weeks, students take another round of CASE21 tests to measure their progress or lack thereof. Teachers log the previous data with the new data for the school and district to evaluate. The goal is growth toward proficient or advanced scores. CASE21 tests provide data to the school district on the projected achievements of the students at the end of the school year. These district assessments prepare schools for the state tests. State tests are the highest form of tests students take. The results of these tests determine school funding, school closure or combinations, and teacher supplies. School districts use their projection tests to prepare administration, teachers, and students for upcoming state tests.

The purpose of an assessment must be decided in order to give the assessment properly, which in turn determines the assessment's validity. Validity is the reason and evidence for giving the assessment based on the accuracy and reliability of assessment choice (Popham 99). Assessment choice ensures the proper evaluation of student progress or mastery within or at the end of a unit. The first criteria in deciding the assessment type are the skills needing to be mastered and the overall demonstration the student must show. The first three Common Core ELA anchor standards “[focus] on the command on evidence... and drawing on evidence as [students] examine the ideas, structure, or style” (Tampio 43), determine

and analyze central idea and theme (43), and “analyze how and why individuals, events, or ideas develop and interact” (43). The last six anchor standards require students to analyze and show evidence supporting the text's craft while integrating ideas (44) by providing “textual evidence to answer questions about words and phrases (4), structure (5), point of view (6), content in diverse media and formats (7), reasoning (8), and similarities to another text (9)” (43). Anchor standards such as one through three require students to identify elements from the text and to answer questions, while standards four through six require students to analyze elements of the text. Standards seven through nine require integration, which is visible in compositions about a text.

Analysis standards include the meaning of the author's choice of language, point of view, structure, flow, and elements employed to develop the story's purpose. Valid assessment scores, whether they show mastery or not, need to be based on an assessment that directly displays student use of the required skills. When asking students to master analysis standards, teachers must choose an assessment that reliably shows students demonstrating an understanding of these concepts. Multiple-choice tests that ask questions about the author's purpose indicate students' mastery of anchor standard four, but a binary choice assessment such as true or false asking students to memorize quotes from a story does not assess mastery. An assessment asking students to create a project highlighting the theme, identifying supporting details that prove the theme, and determining the structure of the text and what type of narrator is evident, reliably assesses student mastery and creates validity in assessment choice. Standards that fit into categories seven through nine require the integration of one through six to create new texts. An assessment choice that best demonstrates content connection among media, reasoning, and comparison is best seen in competitions. Asking students to write on the differences seen in a social media post and a multicultural story demands that a statement be made and that students use evidence and reasoning to defend it. Comparing different works requires a background knowledge of analysis and is seen in a composition assessment. Asking students true and false or multiple-choice questions about the similarities between works gives them the ability to guess, since titles can give away meanings or students can easily assume similarities between works. Composition questions such as “What are the similarities between character A

and character B? Describe how they are similar in their actions and decisions” give students an in-depth question to answer and provides teachers with an accurate assessment of skills.

Composition questions and assignments also assess grammar, formatting, and essay formulation. ELA uses MLA formatting for its longer papers, and college courses require MLA, APA, Chicago, and other format types. When students graduate high school, they need to understand how to follow a format. Grammar is necessary for any area of life because grammar makes conversations, emails, and texts make sense. A sentence’s entire meaning changes without the correct verb tense and subject-verb agreement. Essays require an introduction, a claim, supporting details, and conflicting viewpoint, and must flow in order to convey the purpose and to solidify the impact of the essay. The ability to make a claim, build an argument, use supporting details, find evidence, and assess anchor standards teaches students how to argue, inform, or persuade others in the real world. A conversation is sometimes a spoken argumentative essay, and without proper claims and evidence, the argument is lost. Telling a friend about the previous night’s baseball game is a narrative, and without proper transitions the order of the story is lost.

Standards that include an understanding and command of the English Language while analyzing text to determine the significance of figurative language, connotations, and hidden meanings demands a background knowledge of certain vocabulary. When assessing a student’s vocabulary, there are many approaches teachers take. Writing prompts where using vocabulary words are necessary, binary and multiple-choice quizzes, and short answers are all viable options, and the teacher’s decision must be based on the purpose of assessment. Is the standard requiring denotative meaning of grade-level words? Do students need to interpret hidden meanings? Is the purpose to determine the students’ ability to use new vocabulary correctly? To choose the correct assessment, the proper questions for the outcome must be established. Standards including the interpretation and inference of words or phrases to determine the author’s meaning requires an assessment that asks students to determine the meaning of words in context; therefore, both short answer interpretation or multiple-choice questions benefit this purpose. However, standards where students display command of the English language need an assessment that calls for the use of grade-level vocabulary

in context through a writing prompt or reading response.

Writing prompts and reading responses also expect students to understand what is being read. Comprehension is the ability to understand something, and anchor standards require comprehension of the elements of literature, central ideas, theme, and figurative language. If the teacher’s purpose is to assess the students’ ability to determine theme or central idea and cite textual evidence, then the assessment will need to be a constructed response. However, if a teacher’s goal is to determine whether the students observed details like the color of a door or the number of days the story spanned, then multiple-choice questions are a better option. Measuring students’ mastery of figurative language is case based on what the desired level is. Multiple-choice assessments are helpful for a definition check, while constructed responses or a project are more appropriate when using or identifying examples of figurative language in a text. Identifying story elements, plot, setting, conflict, and resolution are better judged using multiple choice or short answers because the response is brief and does not need an explanation. If students must explain or cite evidence supporting an answer, then constructed response is necessary. Argumentative essays may benefit from a scenario-based assessment where

the scenario poses a controversial issue; the lead-in tasks ask the student to read and summarize arguments given in the source materials, critique those arguments, and analyze the argument; and the culminating essay calls for the students to resent his or her position and reasoning using evidence from the sources. (Zhang et al. 74)

There are various assessment formats teachers must wade through in order to decide how to assess their students in the ELA classroom best. Selected response, critical response, and computer-based assessments all have their place in the ELA classroom; therefore, learning the differences and intentions of each assessment is essential to a teacher’s overall effectiveness. Choosing an assessment format also depends on factors such as time, coherence, and high-stakes language.

Selected response assessments are binary choice, multiple-binary choice, multiple choice, and matching (Popham 155). Binary choice presents students with two options to choose between, most commonly true

and false (162). True and false is time-effective and provides students with a fifty percent chance of answering the question correctly; however, students struggle with negatively-worded binary-choice questions (163). The ELA state tests do not require students to answer true/false questions (Mississippi Department of Education). The level of language in true/false questions needs to be on grade level so students are meeting standards. “Multiple-binary choice is one in which a cluster of items is presented to students, requiring a binary response to each of the items in the cluster” (Popham 164). Presenting students with multiple-binary choice questions gives students practice deciphering content and using that same content to answer multiple questions. State tests require students to read a passage and answer questions regarding the theme; then, the following one refers to the question regarding the theme in the same text. These state questions build on one another. Multiple-binary choice is a way to synthesize the idea of using one piece of information over multiple questions. Multiple-choice assessments present students with one question and three or more possible answers to choose from. State tests, the ACT, and the SAT are all examples of multiple-choice assessments. The reading portion of standardized tests are multiple-choice and all rely on students to use passages to answer correctly (“Student Assessments”). Students use comprehension when reading the passage to answer questions regarding details, use analysis when asked for the author’s intent or to infer information, and use vocabulary to understand the questions and passages. Matching questions require students to match one piece of information to the corresponding answer from a list of options, which measures students’ ability to memorize information.

The second assessment format to discuss is a critical response assessment. Critical response assessments include essays, short answers, fill-in-the-blank, and writing prompts. Essays measure comprehension and ability to use evidence from a passage to create an argument, persuade the reader, or inform an audience. Essays are beneficial when the purpose of the assessment is to measure students’ command of the English language and cite textual evidence. Short answer formats are beneficial for a quick assessment of content knowledge, definitions, or immediate response. Short answer questions compel students to answer concisely and provide an answer based on an opinion of a passage or write a definition

for a vocabulary word. Short answer questions benefit students on later tests because “attempting to retrieve information engages active processing that can enhance later memorability, [and] retrieval tests that engender more effortful, generative processes (e.g., short-answer tests) can enhance later memory more than those completed with relative ease (e.g., multiple-choice)” (McDermott 4). These definitions are either the denotation or connotation of a word in a passage or sentence. Then, fill-in-the-blank questions help students recall definitions when identifying the word the definition describes or answer with a theme or central idea when describing a passage. Spelling, denotative meaning, or identifying connotation are all assessable through fill-in-the-blank questions and are valuable precursors to short answer because they scaffold information, beginning with word identification.

The final assessment format up for discussion is computer-based assessments (CBAs). CBAs provide students with an informational passage and corresponding comprehension questions. The program alerts students if they are wrong or not with a small prompt praising or redirecting their train of thought. If the students answer incorrectly, the application provides hints and feedback for them to reconsider their answers. Computer-based assessments use computer programs to provide immediate feedback to increase reading comprehension. Hooley and Thorpe report that the “timing in which the feedback is delivered also plays an important role in learning outcomes” (1217), because “[i]n line with research on effective formative assessments, the feedback for this study was programmed to be positive, student-friendly, and helpful about where correct answer responses might be found in the reading if students responded wrong” (Hooley and Thorpe 1221). CBAs allow teachers to provide feedback, which in turn allows students to change their thinking on formative assessments. In other words, CBA applications “capitalize on the known motivational benefits of assessments” (1218). Research is still being conducted on the use of CBAs in the classroom, but Hooley’s report says that even if student comprehension does not increase, their confidence does (1232). Computer-based assessments are useful for formative assessments and for helping students modify their way of thinking; however, online applications for summative assessments without feedback

are also useful for teachers and their students. Online assessments are programmed with different formats and are delivered electronically.

When choosing an assessment format, the purpose must be established beforehand, as well as factors such as time effectiveness, curriculum coherence, high-stakes language, and format coherence among classroom and high stakes assessments. Deciding which assessment formats are most beneficial must be done with consideration of the students and their needs, in addition to the abilities of the teacher and his or her curriculum. “Laboratory work suggests that the format of quizzing (i.e., multiple-choice or short-answer) might influence the effectiveness in enhancing later retention, although cross-format benefits are seen” (McDermott et al. 4). Classroom assessments need to be decided by what is most beneficial to the students and conducive to receiving grades on time. Assessing what they learn, how they will be tested statewide, and language coherence are all deciding factors for teachers.

Time effectiveness not only refers to the amount of time the assessment takes, but also the amount of time it takes to grade. Students want feedback from their assessments quickly. The fastest formats are selected response or fill-in-the-blank assessments, which allow for an answer key with no flexibility. Multiple choice and binary choice questions can be graded within a few minutes because the answers are in a numbered list and teachers can outsource grading to teaching assistants because the answers are concrete. Computer-based assessments are “both time saving and staff saving” because the assessment is graded immediately by the application (Hooley and Thorpe 1218). Multiple-choice assessments “optimize” classroom time (McDermott 17); however, multiple-choice assessments or benchmarks “may not necessarily provide a clear signal of what students are learning” (Kaufman et al. 15) because of the probability of guessing answers correctly.

Experiments conducted at Washington University between short-answer and multiple-choice assessments “reveal that low-stakes multiple-choice quizzes with immediate correct-answer feedback can indeed enhance student learning for core course content, as revealed in regular in-class unit exams” (3). Fill-in-the-blank assessments require a little more reading and spell-checking, but the answers are still concrete and require students to answer with one or two words. The only possibility for varying answers is within a first or last name. Teachers

reduce this variance by specifying that students use the last name of the person they are answering with.

Coherence is “the extent to which key system components related to teaching and learning are consistent with each other (or reinforce each other) and all provide the same signals and support to teachers and leaders about what instruction should look like” (Polikoff et al. 1). Coherence among classroom assessments and curriculum, classroom assessment language, high-stakes language, and assessments given or created by teachers must align with standards and supports with the curriculum. Assessments need to focus the instruction on what is important in the curriculum and how to scaffold information for the students. Formative assessments must evaluate lower-level skills that build to the summative assessment skills.

Kaufman reports in a case study that after interviews with administrators and teachers in Louisiana on new standards and assessments, “half of those with whom we spoke said that not only were the mathematics and ELA standards better-aligned to the state assessments than in previous years, but the current state assessments were also more rigorous than previous assessments in a way that held teachers and students to higher expectations” (11). The increased rigor of state assessments means classroom assessments need to align better with high-stakes assessments.

Practicing format coherence among classroom assessments and high-stakes assessments is a consideration for choosing appropriate assessments. Most standardized tests are multiple-choice with an optional writing portion (“Student Assessment”). Students need to practice their high-stakes testing abilities in the classroom so they are not surprised once they reach state tests, the ACT, or the SAT. Having some assessments mimic standardized tests, whether that be practice ACTs or CASE21 tests, allows students to practice their ability to take multiple-choice tests where some questions are dependent on a text or a previous answer. Coherence among assessment and curriculum, language among classroom and high-stakes assessments, and classroom assessment formats and high-stakes assessments create more beneficial assessments for the students.

No assessment format is the most beneficial. Good assessment formats are mutually beneficial to the teacher and his or her students. When deciding which assessment to give to students, whether formative or sum-

mative, time effectiveness and coherence are the most important factors. Furthermore, the most beneficial assessment format varies because each class is different, and every student has different strengths. The average classroom has around 21 students (National Center for Education Statistics) and is comprised of 21 students' strengths, weaknesses, likes, dislikes, temperaments, and more. Humans are complex beings, so the teacher must observe the classroom during formatives, instructional times, and free time to decide which assessment works best for both the teacher and the class.

Different students need different modifications in order to grasp new concepts; therefore, differentiation is implemented to meet the needs of as many students as possible at one time. Differentiation "allows all students to access the same classroom curriculum...designed to meet unique individual needs" (Watts-Taffe 304). In assessment format decisions, teachers must account for the individual needs of each student to assess all of them equitably. Teachers must "[a]ssess students carefully and regularly using a variety of formative assessment tools, then analyze the resulting data to determine patterns of need..." (Watts-Taffe 309). Class size is a major contributor to assessment format decisions. Teachers have different options for class-wide formative assessments, but if the class is too big to handle group discussions in an orderly manner, then a discussion formative would not be effective. If the class size is too big to grade five-page essays promptly, then deciding to have a long essay is not the most beneficial assessment choice. The last difference determining the most beneficial assessment format is the age of the students. The format must be coherent to the standards students must master. For example, kindergarteners will not benefit from the same assessments as high school students. Teachers must consider student age, needs, and class behavior and size when choosing the most beneficial assessments for their students.

There is no overarching assessment format in the English Language classroom. Deciding which assessment format is most beneficial is based on the individual needs of the class and the teacher. The time effectiveness of administering a formative or summative assessment as well as the turnaround of grades and feedback affects assessment choice. Also, the purpose of the assessment influences the assessment format. Students do not need to take a binary-choice assessment when the purpose is to assess their ability to compose an ar-

gumentative essay. Lastly, student differences in learning influence what types of formative assessments are given, alongside what accommodations or materials are given during a summative assessment. Treating students equally in the assessment process is most important.

Assessments are necessary for the classroom. Formative assessments provide students with immediate feedback to change their way of thinking during instruction and show teachers what needs to be modified to meet the students' needs. Summative assessments measure what students master during an instructional unit. Some studies say multiple-choice assessments are not suitable for benchmarks, yet high-stakes tests—like state tests or the ACT—are multiple-choice due to time effectiveness.

There are multiple considerations to account for in the classroom. Teachers need to do their best to meet the needs of their students while effectively assessing them. Learning from formative assessments and modifying instruction before the summative assessments provide teachers with growth data of their students. Teachers will never determine the perfect assessment. The sooner teachers, administrators, and parents realize that teaching and assessments are a process and not a comprehensive determinate of the student or teacher's abilities, the process for ascertaining the most beneficial assessments for the English classroom can begin.

Works Cited

- Hooley, Diana S., and Justin Thorpe. "The Effects of Formative Reading Assessments Closely Linked to Classroom Texts on High School Reading Comprehension." *Educational Technology Research and Development*, vol. 65, no. 5, Springer, 2017, pp. 1215-38, <http://www.jstor.org/stable/4501873>.
- Kaufman, Julia, et al. *Raising the Bar for K-12 Academics: Early Signal on How Louisiana's Education Policy Strategies are Working for Schools, Teacher, and Students*. RAND Corporation, 2019, <http://www.rand.org/t/RR2303>.
- McDermott, K. B., Agarwal, P. K., D'Antonio, L., Roediger, H. L. III, & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology: Applied*, 20(1), 3–21. <https://doi.org/10.1037/xap0000004>.

- Mississippi Department of Education, Mississippi Academic Assessment Program. English II Practice Test. Questar Assessment Inc. & Mississippi Department of Education, 2020, https://ms.nextera.questarai.com/docs/MS_ELA_Practice_Test.pdf.
- “National Teacher and Principal Survey (NTPS).” National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education, https://nces.ed.gov/surveys/ntps/tables/ntps1718_ftable06_t1s.asp.
- Polikoff, Morgan, et al. Exploring Coherence in English Language Arts Instructional Systems in the Common Core Era. Research Report. RR-A279-1. RAND Corporation, RAND Corporation, 1 Jan. 2020, https://www.rand.org/pubs/research_reports/RRA279-1.html.
- Popham, W. James. Classroom Assessment: What Teachers Need to Know, Pearson, 2020.
- “Student Assessment.” The Mississippi Department of Education, 2022, <https://www.mdek12.org/OSA>.
- Tampio, Nicholas. “English Standards, Close Reading, and Testing.” *Common Core: National Education Standards and the Threat to Democracy*. Johns Hopkins University Press, 2018, pp. 40-57.
- Watts-Taffe, Susan, et al. “DIFFERENTIATED INSTRUCTION: Making Informed Teacher Decisions.” *The Reading Teacher*, vol. 66, no. 4, [Wiley, International Reading Association], 2012, pp. 303–14, <http://www.jstor.org/stable/23321311>.
- Zhang, Mo, et al. “Scenario-Based Assessments in Writing: An Experimental Study.” *Educational Assessment*, vol. 24, no. 2, Apr. 2019, pp. 73–90. ERIC, doi:10.1080/10627197.2018.1557515.