

Keshav Anand

Russell/Alexander

American Studies

31 October 2025

A Synthesis of Research on the Efficacy of Narrative Assessment in Secondary Education

Assessment remains a critical component of education, serving as a means to evaluate student progress and mastery over a given subject. Traditionally, quantitative assessments (e.g. multiple-choice tests, standardized exams) have been preferred for their objectivity and ease of grading (Lamiell). Even today, many standardized tests and important exams are primarily quantitative in nature, with American exams such as the SAT and ACT having recently undergone changes to remove or optionalize essay components (McGrath; Semos). While quantitative assessments have their merits, their defined scope limits their ability to capture a student's critical thinking and creativity (Ku). Conversely, narrative assessments (e.g. essays, projects, presentations) with subjective grading ¹ allow students to express understanding in a more holistic and authentic manner (Ku). Recent developments in educational trends due to the rise of Artificial Intelligence (AI) following the COVID-19 pandemic have led many educators to reconsider their methods, accelerating changes in curriculum and assessment (Kamalov et al.). Hence, this paper synthesizes research on the efficacy of subjective narrative assessment in secondary education when compared to traditional quantitative assessments.

A central advantage of narrative assessments is their ability to allow students to demonstrate higher-order thinking skills. While quantitative assessments often reward students for accurate recall of facts, narrative assessments provide a means to distinguish a satisfactory understanding from a nuanced comprehension. This is particularly important to distinguish gifted students; while a standard 4-option multiple-choice question gives full credit for knowing the tested concept (and a 25% change of guessing correctly), both the student who barely understands

¹Structured written responses scored with a non-flexible rubric, such as numerical math problems or essay questions with rigid criteria, are not considered narrative assessments in this context.

a concept and the student who has mastered it receive the same score (Liu et al.). University professors second this notion, agreeing that multiple-choice questions (MCQs) are limited in their ability to assess higher-order cognition (Liu et al.). In the context of a secondary education setting, the impact of narrative assessments are particularly pronounced due to the focus on knowledge retention. Failing to properly distinguish between levels of understanding can inadvertently incentivize surface-level memorization, where students fail to recall basic information learned in a class after the subject is no longer tested (Kooloos et al.). In fact, Kooloos et al. found that students lost about 33% of gained knowledge when learning is geared towards short-term recall. Known as the "forgetting curve," this rapid loss of information can only be mitigated by constant revisiting of material or understanding concepts at a deeper level (Kooloos et al.). Due to the open-ended nature of narrative assessments, students are encouraged to engage with material more deeply, and when educators reward complex understanding, students are incentivized to internalize concepts rather than memorize facts (Ku). Additionally, these benefits are compounded in secondary education as it promotes creativity, which is beneficial for growth and development at younger ages (Redó et al.). When considering that growth and development is the sole focus of secondary education itself, the creativity garnered by narrative assessments is an appealing option for educators.

Another vital component of narrative assessments is their ability to thoroughly evaluate student communication skills. In an increasingly interconnected world, the ability to effectively communicate ideas is paramount, with narrative assessments providing a platform for students to hone these skills (K. J. Wilby et al.). In fact, K. J. Wilby et al. found that narrative assessments even have merits in STEM (Science, Technology, Engineering, and Mathematics) subjects, where they researchers found that narrative assessments have use in the medical industry for summative decision making. Due to the ubiquitous importance of writing and communication skills across disciplines, educators' primary rationale for using narrative assessments lies in their ability to foster these skills (K. Wilby et al.). These motives are backed by scientific evidence beyond communication skills, as it has been repeatedly proven that the ability to communicate a certain

concept — especially when supported by metacognitive strategies such as planning, monitoring, and evaluation — is correlated with mastery of that concept (Hamzah et al.). While narrative assessments promote such communication skills, students who are not proficient writers may be disadvantaged. Lo et al. found that assessing students through essays requires an implied mastery of the English language, which may unfairly penalize non-native speakers. This issue is particularly sensitive in secondary education, as the vast percentage of immigrants and non-native speakers will have their course evaluation depend on their English proficiency. Hence, while narrative assessments promote communication skills, educators must be wary of potential biases against non-native speakers.

While proponents of narrative assessments highlight their benefits, critics focus on the subjectivity of such assessments as a major flaw. Unsurprisingly, bias and subjectivity is a primary complaint of students, where a majority often perceive essays and projects as unfairly graded (Bullock et al.). These complaints are not without merit, as inconsistencies are bound in subjective evaluation due to educator bias and differences between sections and educators. For example, the same history course can be taught by multiple teachers, where each teacher may have different expectations for essay responses. In fact, Lopera-Oquendo et al. found that inter-rater reliability (IRR) for essay grading was often low, with even rubric-based grading systems failing to ensure consistency between graders. This inconsistency can lead to student frustration and a perception of unfairness, which can negatively impact student motivation and engagement. While many focus on the negative aspects of subjectivity, some researchers argue that subjectivity can be beneficial in certain contexts. For example, positive expectation bias — where educators hold higher expectations for certain students — can lead to improved performance due to positive reinforcement of a student's potential and abilities (Boer et al.). Positive expectation bias is also useful in flattening performance outliers; for example, a student having a bad day during an assessment date may be unfairly penalized in a quantitative assessment, while positive expectation bias allows teachers to understand and compensate for performative outliers. This not only benefits gifted students but also generally weaker students,

and studies show that teachers sometimes subconsciously use shifting standards to leniently grade struggling (yet hardworking) students (Gil-Hernández et al.).

Works Cited

- Boer, H., et al. “Sustainability of teacher expectation bias effects on long-term student performance.” *Journal of Educational Psychology*, vol. 102, 2010, pp. 168–79.
<https://doi.org/10.1037/a0017289>.
- Bullock, Justin L., et al. “In Pursuit of Honors: A Multi-Institutional Study of Students’ Perceptions of Clerkship Evaluation and Grading.” *Academic medicine : journal of the Association of American Medical Colleges*, 2019.
<https://doi.org/10.1097/acm.0000000000002905>.
- Gil-Hernández, Carlos J., et al. “Teacher Bias in Assessments by Student Ascribed Status: A Factorial Experiment on Discrimination in Education”. *Sociological Science*, vol. 11, no. 27, 2024, pp. 743–76. <https://doi.org/10.15195/v11.a27>.
- Hamzah, Hainora, et al. “Systematic Literature Review on the Elements of Metacognition-Based Higher Order Thinking Skills (HOTS) Teaching and Learning Modules”. *Sustainability*, 2022.
<https://doi.org/10.3390/su14020813>.
- Kamalov, Firuz, et al. “New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution”. *Sustainability*, 2023. <https://doi.org/10.3390/su151612451>.
- Kooloos, J., et al. “The Effect of Passive and Active Education Methods Applied in Repetition Activities on the Retention of Anatomical Knowledge”. *Anatomical Sciences Education*, vol. 13, 2019, pp. 458–66. <https://doi.org/10.1002/ase.1924>.
- Ku, Kelly Y.L. “Assessing students’ critical thinking performance: Urging for measurements using multi-response format”. *Thinking Skills and Creativity*, vol. 4, no. 1, 2009, pp. 70–76.
<https://doi.org/https://doi.org/10.1016/j.tsc.2009.02.001>.
- Lamiell, James T. “Some Historical Perspective on the Marginalization of Qualitative Methods Within Mainstream Scientific Psychology”. *Situating Qualitative Methods in Psychological Science*, Routledge, July 2018, pp. 11–26, <https://doi.org/10.4324/9781351136426-2>.

- Liu, Qian, et al. "Multiple-choice questions (MCQs) for higher-order cognition: Perspectives of university teachers". *Innovations in Education and Teaching International*, vol. 61, 2023, pp. 802–14. <https://doi.org/10.1080/14703297.2023.2222715>.
- Lo, Y., et al. "Assessing content knowledge through L2: mediating role of language of testing on students' performance". *Journal of Multilingual and Multicultural Development*, vol. 44, 2021, pp. 1013–28. <https://doi.org/10.1080/01434632.2020.1854274>.
- Lopera-Oquendo, Carolina, et al. "Rating writing: Comparison of holistic and analytic grading approaches in pre-service teachers". *Learning and Instruction*, vol. 94, 2024, p. 101992. <https://doi.org/https://doi.org/10.1016/j.learninstruc.2024.101992>.
- McGrath, Steve. "College Board Updates on the SAT Essay and Subject Tests", 2021, Accessed: 2025-10-26, info.methodlearning.com/blog/college-board-updates-on-the-sat-essay-and-subject-tests.
- Redó, Núria Arís, et al. "Dimensions of Creativity in Secondary School High-Ability Students". *European Journal of Investigation in Health, Psychology and Education*, vol. 11, 2021, pp. 953–61. <https://doi.org/10.3390/ejihpe11030070>.
- Semos, Kristina. "What's Staying the Same on the New ACT? (Pt 2)", Sept. 2024, Accessed: 2025-10-26, www.ivyloungetestprep.com/blog/act-changes.
- Wilby, K., et al. "Discriminating Features of Narrative Evaluations of Communication Skills During an OSCE". *Teaching and Learning in Medicine*, vol. 31, 2019, pp. 298–306. <https://doi.org/10.1080/10401334.2018.1529570>.
- Wilby, Kyle John, et al. "Reliability of narrative assessment data on communication skills in a summative OSCE". *Patient Education and Counseling*, vol. 102, no. 6, 2019, pp. 1164–69. <https://doi.org/https://doi.org/10.1016/j.pec.2019.01.018>.