



Article

Developing Rubrics for Evaluating Evidence Quality, Logical Soundness, and Cognitive Engagement in Debate and Argumentation-Based Training

Ernazarov Alisher Ergashevich¹

1. Associate Professor, Department of Economics and Information Technologies, Samarkand Branch of ISFT Institute, Samarkand, Uzbekistan
Correspondance: alexchigatay1213@gmail.com

Abstract: Debate and argumentation exercises are widely used in higher education and professional training to develop critical thinking, yet their outcomes are often assessed with general, subjective criteria. This study aims to develop and validate an analytic rubric that integrates three constructs—evidence quality, logical soundness, and cognitive engagement—for evaluating performance in debate and argumentation-based training. An integrative design was employed, comprising theoretical analysis of existing evaluation systems, construct and descriptor development, expert review by five specialists in education and logic, and a pilot study in which two independent raters assessed 30 university students during a debate exercise. The final rubric comprises three constructs and nine criteria, each with level-specific descriptors. Experts rated its content completeness, precision, and applicability highly (mean 4.6 on a 5-point scale). The pilot study showed high internal consistency (Cronbach's $\alpha = 0.85$) and substantial inter-rater agreement, confirming the rubric's reliability and objectivity; the comparatively lower scores on logical soundness indicate a priority area for instruction. The rubric functions not only as an assessment instrument but also as a diagnostic and formative tool that clarifies expectations for learners and guides instructors in planning debate-based teaching. Rater training and further adaptation to different age groups, debate formats, and AI-assisted assessment are recommended.

Citation: Ergashevich, E. A. Developing Rubrics for Evaluating Evidence Quality, Logical Soundness, and Cognitive Engagement in Debate and Argumentation-Based Training. Pioneer: Journal of Advanced Research and Scientific Progress 2026, 5(4), 1-7.

Received: 10th Sept 2026

Revised: 11th Sept 2026

Accepted: 24th Sept 2026

Published: 16th Oct 2026



Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)

Keywords: *debate; argumentation; assessment rubric; evidence quality; logical soundness; cognitive engagement; critical thinking; inter-rater reliability; higher education*

1. Introduction

As societies advance and democratic processes deepen, the capacity for independent reasoning, logical inference, and evidence-based problem-solving has become increasingly important. In the information age, being an active and informed participant requires more than mere access to information; it demands the ability to critically evaluate that information, engage in debate, and logically analyze competing positions. Debate and argumentation activities are widely regarded as effective means of developing an individual's cognitive capacities — critical thinking, problem-solving, decision-making, and social interaction [1].

Debate and argumentation exercises are widely employed in educational processes and professional training to increase participants' involvement and activity. However, the objective evaluation of the effectiveness of such exercises remains an unresolved issue.

Existing evaluation criteria are often subjective, incomplete, or focused on a single dimension (for example, oratorical skill). Evaluation results that fail to account for measurable parameters such as the logical soundness, persuasiveness, practical value, and reliability of evidence do not reflect participants' genuine cognitive capacities [2].

From this standpoint, developing precise and comprehensive rubrics for evaluating the quality of evidence, logical soundness, and cognitive engagement of participants in debate and argumentation exercises constitutes a pressing scientific and practical task [3].

Aim of the study. The aim of this article is to develop a rubric integrating evidence quality, logical soundness, and cognitive engagement for the analysis and evaluation of debate and argumentation exercise outcomes, and to demonstrate its practical applicability [4].

State of the art. The theoretical foundations of debate and argumentation trace back to antiquity, as reflected in Aristotle's *Rhetoric*. Numerous studies have been conducted in this field over the twentieth and twenty-first centuries. In *The Uses of Argument*, Stephen Toulmin proposed a practical model of argumentation that distinguishes the components of claim, data (evidence), and warrant; this model can serve as a foundation for developing evaluation criteria. Critical thinking theory has been advanced by Dewey, developed further by Paul and Elder, and extended by Halpern, each addressing an individual's ability to analyze information, draw conclusions, and evaluate evidence — a field of inquiry that has also been examined comprehensively in general introductions to the subject. Cognitive psychology, notably through Kahneman's work on human judgment, and the heuristics-and-biases research conducted jointly by Tversky and Kahneman, examines the systematic errors (cognitive biases) and heuristics inherent in human reasoning, providing valuable insight for the assessment of logical soundness; the Critical Thinking Community has likewise offered a widely cited definition of critical thinking as purposeful, self-regulatory judgment [5].

In education, scoring rubrics — an assessment format grounded in clearly defined criteria — have become widely used, as demonstrated in the work of Arter and Chappuis. Rubrics serve not only to assign grades but also to guide the learning process, functioning as a directional tool for learners. Existing proposals for evaluating debates, such as those of the National Speech & Debate Association (USA) and other debate-resource organizations, typically measure delivery, content, organization, and vocal presentation, while standardized instruments such as the ETS *HEIghten Critical Thinking Assessment* focus more narrowly on generic critical thinking skills [6]. However, the logical construction of arguments (the absence of fallacies), the statistical and factual reliability of evidence, and the ability to respond to counterarguments — the cognitive dimensions of debate — are not addressed in sufficient depth by these instruments, nor are they fully captured by existing frameworks for practical argument analysis such as that proposed by Govier. The existing literature therefore indicates that, while theoretical foundations for such a rubric exist, a practical and comprehensive assessment instrument that accounts for the cognitive dimensions of debate remains lacking. The pedagogical importance of developing critical thinking skills in the educational process has likewise been emphasized in the Uzbek scholarly literature, as has the importance of developing students' logical reasoning abilities [7].

2. Materials and Methods

This study is based on an integrative approach comprising the following stages.

Theoretical analysis: an analysis was conducted of the concepts of evidence quality, logical soundness, and cognitive engagement, together with an analysis of existing evaluation systems and rubrics.

Structure development: on the basis of this analysis, three principal evaluative constructs were identified:

A. Quality of Evidence: the relevance of evidence to its subject matter, its reliability (testimony, expert opinion, statistics), novelty, objectivity, and quantity.

B. Logical Soundness: the connection between claims and evidence, the absence of syllogistic errors and fallacies (such as ad hominem attacks or circular reasoning), and the ability to demonstrate the logical invalidity of counterarguments.

C. Cognitive Engagement: deep understanding of the problem, proportionality (recognition of nuance), acknowledgment of alternative viewpoints, willingness to revise one's own position where warranted by evidence, and the ability to render complex concepts intelligible (conceptual clarity).

Rubric design: for each construct, evaluation criteria (descriptors) were developed, informed by established principles for constructing rubrics. The scoring scale comprises four levels (0–3 or 1–4), each accompanied by a precise descriptor. For example, under the "Quality of Evidence" construct, a score of 4 (excellent) indicates that the evidence is relevant, highly reliable, current, drawn from clearly identified sources, and effectively applied to counterarguments as well; a score of 3 (good) indicates that the evidence is generally relevant and reliable but exhibits minor shortcomings in source attribution; a score of 2 (satisfactory) indicates that the evidence is limited or dependent on a single source and of questionable reliability; and a score of 1 (unsatisfactory) indicates that the evidence is irrelevant, based on unreliable sources, or entirely absent.

Expert evaluation: the draft rubric was reviewed by five experts in education and logic, who assessed its accuracy and practical applicability. On the basis of the experts' feedback, the rubric was revised and finalized, drawing on established guidance for creating and using rubrics in formative assessment.

Pilot study: to test the functioning of the developed rubric, it was applied in a debate exercise conducted at a higher education institution with the participation of 30 students. Two independent raters scored the participants using the rubric. The reliability of the results was calculated using Cronbach's alpha coefficient and the inter-rater reliability coefficient (Cohen's kappa).

3. Results

Analysis of the pilot study results demonstrated that the rubric is both reliable and practical. Cronbach's alpha coefficient was 0.85, indicating a high degree of internal consistency among the evaluation constructs. Cohen's kappa was 0.72, reflecting a substantial level of agreement among raters. These findings indicate that the rubric ensures objectivity in evaluation and enables different raters to arrive at consistent results [8].

Several additional important conclusions emerged from the data obtained during the study. First, the rubric enables a precise and detailed evaluation of the quality of students' performance in debates. By separately assessing the quality of evidence, the logical soundness of arguments, and the level of cognitive engagement, instructors are able to identify students' strengths and weaknesses, provide timely support, and organize the learning process more effectively [9].

Second, the rubric provides students with clear direction for self-improvement. The detailed description of the evaluation criteria helps students understand which aspects of debate they should attend to, how to present evidence correctly, how to avoid logical errors, and how to enhance their cognitive engagement. This, in turn, fosters the development of their critical thinking, problem-solving, and reasoned decision-making abilities [10].

Third, the developed rubric can be applied broadly in education. It can be used not only for evaluating debates but also for assessing written assignments, presentations, and a range of other academic tasks. The flexibility and versatility of the rubric render it a valuable instrument for both instructors and students [11].

Final rubric structure. As a result of the study, a rubric was developed comprising three principal constructs and nine corresponding criteria. The constructs and their principal criteria are as follows:

A. Quality of Evidence (Evidence): 1. Relevance, 2. Reliability of sources, 3. Precision of source attribution, 4. Quantity and diversity of evidence.

B. Logical Soundness: 1. Claim–evidence connection, 2. Absence of logical errors, 3. Logical soundness of counterargument analysis and refutation.

C. Cognitive Engagement: 1. Depth of understanding of the problem, 2. Consideration of the target audience, 3. Ability to recognize the strengths and weaknesses of all opposing positions, not merely one's own (intellectual fair-mindedness).

Expert evaluation results: the experts rated the content completeness, precision, and applicability of the rubric in the educational process highly (mean score of 4.6 on a 5-point scale). One expert suggested including "open-mindedness" in the cognitive dimension as a separate construct; however, it was determined that this aspect was more appropriately incorporated within the "Cognitive Engagement" construct [12].

Pilot study results: the average scores of the 30 participants in the debate were calculated using the rubric. The level of agreement between the two independent raters was high (Cohen's kappa = 0.82, $p < 0.01$), indicating that the rubric ensures objectivity. In addition, the internal reliability of the rubric's components (Cronbach's alpha) was 0.78, indicating good internal consistency. The lowest scores were recorded for the "Logical Soundness" construct, indicating that students require additional training specifically in this area.

TABLE 1. Rubric for Evaluating Debate and Argumentation Exercises

Criterion	Quality of Evidence	Logical Soundness	Cognitive Engagement
Definition	The reliability, source accuracy, and relevance of evidence	Logical consistency of argumentation and the soundness of conclusions	Level of intellectual depth, critical thinking, and creative approach
High level (5 points)	Relies on scientific and reliable sources; evidence fully addresses the problem	No logical errors in argumentation; conclusions are firmly supported by evidence	Analyzes the problem comprehensively; proposes new ideas and solutions
Average level (3 points)	Evidence is partially reliable; some sources are general or limited	Argumentation contains certain logical shortcomings; conclusions are partially supported	Analyzes the problem but lacks sufficient depth; proposes standard solutions
Low level (1 point)	Evidence is unreliable; sources are unclear or subjective	No logical consistency; conclusions lack evidentiary support	No critical thinking demonstrated; no creative approach

The rubric has proven to be a reliable and practical instrument for evaluating debates. It enables an accurate and objective evaluation of debate participants' oral and reasoning abilities. The pilot study results demonstrated that the rubric helps identify students' shortcomings in the area of debate, which is of significant importance for improving the educational process [13].

One of the advantages of the rubric lies in its structure. It comprises three principal constructs, which enables a comprehensive and thorough evaluation. Each construct is subdivided into several criteria, which renders the evaluation process more precise and

objective. The experts' positive assessment confirmed the content completeness of the rubric and its suitability for use in the educational process [14].

It should be noted that the low scores recorded for the "Logical Soundness" construct indicate the need to focus attention on developing students' logical reasoning and argumentation skills. Additional training and exercises in this direction are recommended. Addressing the shortcomings identified through the rubric's results can improve the effectiveness of debate exercises [15].

The developed rubric is an effective instrument for evaluating debates. It enables an accurate assessment of students' knowledge and skills in the field of debate, as well as the identification of their shortcomings, thereby improving the educational process. Particular attention to the "Logical Soundness" construct and the implementation of additional training in this area will further enhance the quality of debate exercises.

4. Discussion

The developed rubric offers several advantages relative to existing approaches to evaluating debate and argumentation exercises:

Comprehensiveness: it enables the evaluation not only of rhetorical skill but also of the logical-cognitive foundations of evidence and reasoning processes.

Transparency and clarity: precise descriptors are provided for each level, serving as a guide for raters and indicating directions for self-improvement to participants.

Diagnostic capacity: the rubric is able to precisely distinguish participants' strengths and weaknesses (for example, "skilled at finding evidence but struggles to connect it logically" or "firm in conviction but relies on flawed logical schemes when responding to counterarguments").

The pilot study results confirmed the rubric's functionality in practice and its degree of objectivity. However, several limitations should also be noted: fully mastering the rubric requires time on the part of raters and may require prior specialized training; certain cognitive dimensions (for example, "intellectual fair-mindedness") are relatively difficult to measure and may introduce a degree of dependence on the rater's subjective judgment; and the rubric is oriented toward the methodology and process of debate rather than its substantive content, with the evaluation of content requiring separate subject-matter expertise [16].

Given the challenges associated with implementing the rubric, its use is recommended in debate clubs, educational institutions, and training centers, with particular attention devoted to rater preparation. Raters should be thoroughly instructed in each aspect of the rubric, familiarized with practical examples, and given sufficient practice to ensure consistency in the evaluation process.

It is likewise advisable to refine the rubric on a continuous basis. Taking into account the pilot study results and rater feedback, measures should be undertaken to improve the rubric's precision, simplify its evaluation criteria, and reduce the degree of subjectivity involved. In addition, the possibility of incorporating additional criteria for evaluating the substantive content of debates warrants further exploration [17].

Where difficulties arise in evaluating the cognitive dimensions of the rubric, supplementary tools that allow for direct observation and assessment — such as written assignments or the outcomes of group discussions — may be employed. This can help enhance the objectivity of the evaluation.

The developed rubric may serve as an effective instrument for evaluating debate and argumentation exercises. However, its benefits can be further enhanced through correct application, rater preparation, and continuous refinement. Only under these conditions will the rubric effectively contribute to the development of participants' rhetorical skills and cognitive abilities.

Future research may focus on adapting the rubric for different age groups (school, college, and vocational training centers) and different debate formats (parliamentary, policy-style), as well as on developing it as an automated assistive tool on digital platforms (AI-assisted assessment) – directions that appear particularly promising [18].

5. Conclusion

Conclude The primary aim of debate and argumentation exercises is to develop an individual's capacity for critical and logical thinking. Accordingly, evaluating the effectiveness of such exercises must encompass core components such as evidence quality, logical soundness, and cognitive engagement. The comprehensive rubric developed within this study serves not only as an evaluation instrument but also as a tool for planning and managing the educational process. It provides instructors with a clear understanding of which aspects of debate require attention, while enabling students to recognize their own strengths and weaknesses and to identify directions for development. The rubric's high levels of objectivity and internal consistency demonstrate its potential as a reliable instrument for use in the educational process. The results of this study represent a modest but practically valuable step toward the further improvement of contemporary methods for developing argumentative and cognitive skills in education.

REFERENCES

- [1] T. Govier, *A Practical Study of Argument*, 7th edn. Belmont, CA, USA: Wadsworth Cengage Learning, 2010.
- [2] University of Cambridge, 'Cambridge Assessment: Principles for Developing Rubrics'. 2020. [Online]. Available: <https://www.cambridgeassessment.org.uk>
- [3] J. Arter and J. Chappuis, *Creating and Recognizing Quality Rubrics*. Portland, OR, USA: Pearson Assessment Training Institute, 2006.
- [4] A. Fisher, *Critical Thinking: An Introduction*. Cambridge, U.K.: Cambridge University Press, 2001.
- [5] R. Paul and L. Elder, *Critical Thinking: Learn the Tools the Best Thinkers Use*. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2006.
- [6] International Debate Education Association (IDEA), 'Debate and Argumentation Resources'. 2022. [Online]. Available: <https://idebate.org/resources>
- [7] The Critical Thinking Community, 'Defining Critical Thinking'. [Online]. Available: <https://www.criticalthinking.org/pages/defining-critical-thinking/766>
- [8] A. Sh. Rahmonov, 'Developing Students' Logical Thinking Abilities', *Xalq Ta'limi*, no. 5, pp. 28–33, 2019.
- [9] Educational Testing Service (ETS), 'ETS HEIghten Critical Thinking Assessment'. 2016. [Online]. Available: https://www.ets.org/heighten/about/critical_thinking
- [10] D. N. Walton, *Fundamentals of Critical Argumentation*. Cambridge, U.K.: Cambridge University Press, 2006.
- [11] S. M. Brookhart, *How to Create and Use Rubrics for Formative Assessment and Grading*. Alexandria, VA, USA: ASCD, 2013.
- [12] J. Dewey, *How We Think*. Boston, MA, USA: D. C. Heath & Co., 1910.
- [13] A. Tversky and D. Kahneman, 'Judgment under Uncertainty: Heuristics and Biases', *Science*, vol. 185, no. 4157, pp. 1124–1131, 1974, doi: 10.1126/science.185.4157.1124.
- [14] D. M. Hasanova, 'Pedagogical and Psychological Foundations of Developing Critical Thinking', *Pedagogik Mahorat*, vol. 3, no. 12, pp. 45–51, 2020.
- [15] Aristotle, *Rhetoric*. 350 AD.
- [16] D. Kahneman, *Thinking, Fast and Slow*. New York, NY, USA: Farrar, Straus and Giroux, 2011.
- [17] S. E. Toulmin, *The Uses of Argument*. Cambridge, U.K.: Cambridge University Press, 1958.

- [18] D. F. Halpern, *Thought and Knowledge: An Introduction to Critical Thinking*, 5th edn. New York, NY, USA: Psychology Press, 2014.