

Exploration of the Concept of “Evidence-Based Evaluation” in Classroom Teaching Evaluation

—Taking “Literature Reading and Writing” Task Group Practice as an Example

Tengda Qi

College of Arts, Beijing Normal University, Beijing, China

Email: 1009730951@qq.com

How to cite this paper: Qi, T. D. (2023). Exploration of the Concept of “Evidence-Based Evaluation” in Classroom Teaching Evaluation. *Open Journal of Social Sciences*, 11, 640-648.
<https://doi.org/10.4236/jss.2023.1110037>

Received: July 23, 2023

Accepted: October 28, 2023

Published: October 31, 2023

Abstract

The concept of “evidence-based evaluation” is a driving force in the development of the “teaching-learning-assessment” cycle. Based on the typical evidence obtained and structured, it can constitute a comprehensive depiction of the complete learning “person”. The combination of quantitative and qualitative perspectives provides a complete analysis of students’ learning abilities and performance; the use of evidence to reflect on teaching can promote the simultaneous improvement of teachers’ teaching abilities and students’ learning abilities, and drive continuous development of teaching improvement.

Keywords

Obtaining Evidence, Organizing Evidence, Analyzing Evidence, Using Evidence

1. Introduction

“Evidence-based evaluation” is a new assessment concept based on evidence-based theory, which focuses on assessing the level of students’ core literacy development on the basis of extensive collection of evidence of students’ learning outcomes and behavioral process performance. Language learning is characterized by “perception”, “appreciation” and “perception”, which makes teaching evaluation lack of objective evidence support and to a certain extent subjective and arbitrary, making it difficult to effectively reflect on teaching. “Teaching-learning-evaluation integration” seeks to maximize the contribution of evaluation to teaching (Zhong, 2012). “Evidence-based evaluation” can effectively improve the scientificity and relevance of classroom teaching evaluation because of its expressive form of evidence, wide range of evidence sources and personalized organization

of evidence. Evidence about students' key competencies is needed to support classroom assessment (Feng, 2012). Evidence-based evaluation in teaching involves the following steps: obtaining typical cognitive and non-cognitive evidence; analyzing the evidence in a quantitative and qualitative manner and establishing causal associations; and using the evaluation results to reflect on teaching abilities and behaviors. This study draws on the "evidence centered design" (ECD) model of assessment design, and takes the teaching practice of "reading and writing literature" as an example in an attempt to promote the development of the "teaching-learning-assessment" cycle under the guidance of the "evidence-based concept".

2. Obtaining and Organizing Evidence: Cognitive Performance Evidence and Non-Cognitive Evidence Complement Each Other

The application of "evidence-based concept" to teaching evaluation can significantly bring into play the backwash effect of evaluation on teaching and promote the integration of "teaching-learning-evaluation". The key to implementing the "evidence-based evaluation" concept in teaching is to organize, identify, discover, and expand the evidence of student learning, and to structure the evidence to portray the typical performance of students based on high-quality evidence.

2.1. Designing Learning Tasks and Collecting Cognitive Performance Evidence of Students

Evidence of learning comes from a wide range of sources, and "evidence-based evaluation" requires at high quantity and quality of cognitive performance evidence. Cognitive expressive evidence refers to students' language learning ability, which is quantitatively describable learning performances that can be effectively extracted from process products such as recitation, speech, writing demonstration, book exchange, drama performance, and survey interview. Based on the characteristics of the learning task clusters, "determine a complete task model based on specific grade levels and specific topics, select task contexts, and design learning tasks that stimulate student performance" (2017), students generate learning products through oral and written expressions in the classroom driven by the tasks, and teachers focus on students' understanding of the elements of typical competency performance for the task cluster in the learning products, which is an effective path to collect evidence of students' cognitive expressiveness.

2.2. Focusing on Learning Participation and Obtaining Non-Cognitive Evidence of Students

Non-cognitive evidence is also an important source of evidence for "evidence-based evaluation". Non-cognitive evidence refers to students' engagement in learning activities, including cognitive engagement, affective engagement, behavioral engagement, and other key dimensions. Based on the characteristics of this study, cognitive engagement can be defined as a high degree of involvement

in the cognitive strategies and psychological resources used by students in the classroom learning context (Ravindran, Gerun, & Dilbeck, 2005); affective engagement refers to positive emotional responses toward the learning task or others (e.g., teachers and peers) (Ni & Wu, 2011); and behavioral engagement refers to a high degree of involvement in students' participation in the learning process. In task group instruction (Ni & Wu, 2011), teachers should pay attention to extracting the overt behaviors that characterize elements such as students' learning time, active learning, classroom participation, degree of self-monitoring, interest in learning, learning value, and sense of identity to obtain evidence of students' non-cognitive nature.

2.3. Structuring the Elements of Typical Student Learning Ability and Organizing Learning Evidence

Obtaining typical evidence is the key to evaluating the development level of students' core literacy. "Evidence-based evaluation" requires the selection of critical and effective typical evidence. Typical evidence is obtained by relying on curriculum standards that guide the learning content and pedagogical recommendations, as well as by portraying the level of academic quality. In order to have a precise meaning, the dispersed variables should be structured to represent their relationships and generate typical evidence for evaluation of teaching and learning (Feng, 2012). Thus, cognitive and non-cognitive performance evidence should be combined to form a comprehensive picture of the "person" who is learning in its entirety. In addition, local characteristics and curricular differences may result in typical evidence of learning existing outside the formulation of the standards, which needs to be analyzed on a problem-specific basis.

3. Portraying and Analyzing Evidence: Organic Combination of Quantitative and Qualitative Analysis

Establishing rules of evidence and analyzing quantitatively based on evidence can clearly portray the level of students' ability and performance, but "it is difficult to balance the integrity and developability, and to fulfill the developmental function of evaluation. Descriptive-oriented qualitative evaluation is characterized by listening and observation, understanding and interpretation, and researching and improving, which can effectively compensate for the shortcomings of quantitative-oriented classroom evaluation and provide strategies for optimizing classroom teaching" (Zhong, 2012). Therefore, the organic combination of quantitative analysis and qualitative analysis can provide a comprehensive analysis of students' learning ability.

3.1. Establishing Rules of Evidence and Quantitatively Analyzing Evidence of Students' Cognitive Performance

Constructing specific observation dimensions, describing the assigned information, and building rules of evidence are the basis for quantitative analysis of students' typical ability performance (Feng, 2012). Taking the novel units in

the second volume of the compulsory textbook as an example, focusing on the typical ability elements of students' novel reading, centering on the two learning activities of "reading" and "creating", combining Bloom's taxonomy of educational objectives, refining four dimensions of analysis: "comprehension", "analysis", "synthesis", and "application", and establishing rules of evidence to characterize students' typical ability performance. The details are shown in **Table 1**.

According to the rules of evidence, typical evidence can be analyzed quantitatively.

Table 1. Evidence rules for Classroom Teaching Evaluation of the novel units in the second volume of the compulsory textbook.

Indicator		Score 9 - 10	7 - 8	5 - 6	3 - 4	1 - 2
Reading and Creative Writing	Characters	Character image can be precisely portrayed and summarized, the complexities and changes of human nature can be understood and expressed in combination with setting and plot	Character image can be accurately portrayed and summarized, the complexities or changes of human nature can be understood in combination with setting and plot	Salient feature of characters can be summarized, the changes of human nature can be understood in combination with setting and plot	Character image can be simply summarized, the changes of human nature can be understood in combination with setting and plot	Character image can be simply summarized, the complexities or changes of human nature cannot be understood in combination with setting and plot
	plot	The causal logic between plots can be clearly presented in oral expression and composition	The causal logic between plots can be presented in oral expression and composition	The causal logic can be presented in oral expression or composition	The causal logic can be simply presented in oral expression	The causal logic cannot be presented
	setting	Characteristics of the setting can be precisely summarized, and the setting fitting the elements of character, plot, and theme can be shaped	Characteristics of the setting can be accurately summarized, and the relationship between setting and elements of character, plot, and theme can be understood	Salient feature of the setting characteristics can be summarized, and the relationship between setting and elements of character, plot, and theme can be understood	Characteristics of the setting can be simply summarized, and the superficial relationship between setting and elements of character, plot, and theme can be understood	Characteristics of the setting cannot be summarized, and the relationship between setting and elements of character, plot, and theme cannot be understood
	Theme	The theme of the novel can not only be summarized and analyzed in depth, but also skillfully expressed through the elements of character, plot, and setting	The theme of the novel can not only be rationally summarized and analyzed, but also expressed through the elements of character, plot, and setting	The theme of the novel can be rationally summarized and analyzed or expressed through characters, plot, setting and other elements	The theme of the novel can be simply summarized and analyzed, and can also be simply expressed through the elements of character, plot, and setting	The theme of the novel cannot be summarized and analyzed, nor can it be expressed through characters, plot, setting and other elements

Continued

structure	The structural features of the novel can be understood in depth, and appropriate structures are presented in the composition through the selection of narrative perspectives, the shaping of narrative scenes, and the use of appropriate narrative techniques	The structural features of the novel can be understood, and generally appropriate structure can be presented in the composition through the use of appropriate narrative techniques	The structural features of the novel can be understood, or generally appropriate structure can be presented in the composition through the use of appropriate narrative techniques	The structure of the novel can be understood at a superficial level, or the structure with certain features can be presented in the composition through the use of simple narrative techniques	The structural features of the novel cannot be understood, and a generally appropriate structure cannot be presented in the composition through the use of appropriate narrative techniques
language	The subtleties of the novel's language can be deeply appreciated, the words of the characters in the composition correspond to their identities and personalities, and the syntax and rhetoric are rich and relevant	The subtleties of the novel's language can be appreciated, the words of the characters in the composition correspond to their identities and personalities, and the syntax and rhetoric are rich	The subtleties of the novel's language can be understood, or the words of the characters in the composition correspond to their identities and personalities	The subtleties of the novel's language can be generally understood, or the words of the characters in the composition do not correspond to their identities and personalities	The subtleties of the novel's language can be understood

Taking "plot" as an example, some students summarized the plot of "Lin Chong Shelters from the Snowstorm in the mountain Spirit Temple" as "Lin Chong met his old friend Li Xiaoer, learned that Captain Lu Qian and others came to Cangzhou to buy a machete to seek revenge, but was unsuccessful and then slackened. Then he was assigned to the fodder depot, but because of the purchase of wine, he escaped from the fire in the fodder depot, and overheard the plot of Captain Lu Qian, Fu An and others, so he killed several people to take revenge and then went up to the MountLiangshan on a snowy night". The evidence presents the understanding of causal connection between "buying wine" and "escaping the fire", but lacks insight into the implications of "wind and snow" and why the fire was started in the wind and snow. Therefore, this evidence indicates that the student "can present the causal logic of the plot in simple oral expressions" and the plot indicator can be assigned a score of "3 - 4". If the student can pay attention to the connection between Lin Chong's "searching for three or five days", "on the sixth day, Guan Ying calling Lin Chong to the hall of inspection" and "the snow was falling tightly", then it belongs to the category of "being able to present the cause and effect logic between the plots clearly in oral expression", and if the student is able to further develop the causal logic of the plot in his or her composition, the plot indicator can be assigned a score

of “9 - 10”.

3.2. Developing Qualitative Evaluation Tools to Characterize Students' Non-Cognitive Evidence

Based on the learning engagement theory and the level of “non-cognitive factors” portrayed in the curriculum standard, developing qualitative evaluation tools can qualitatively analyze the typical learning performance typical learning performance of students. The details are shown in **Table 2**.

For the secondary dimensions of the framework, assignment descriptions can be used to analyze students' noncognitive evidence. In practice, three levels of analysis can be established in combination with specific learning content. Taking the novel units in the second volume of the compulsory textbook as an example, the deep cognitive strategies for students to summarize the character image of “Lin Chong” can be divided into “analyzing the character’s personalities characteristics through the causal factors of the plot”, “refining character’s emotions through the description of the setting”, and “summarizing characterization through internal and external details of the character”. The students’ deep cognitive strategies for summarizing the characterization of “Lin Chong” can be divided into three levels of analysis, such as “analyzing character characteristics through plot cause and effect”, “refining character emotional characteristics through environmental description”, and “summarizing character image through internal and external details”. By extracting the non-cognitive evidence from a student’s performance in the task of “expressing how he or she summarizes the character image”, the results of the qualitative analysis of the student’s classroom performance can be generated.

Table 2. Analysis framework of “non-cognitive evidence”.

	Primary dimension	Secondary dimension	
Non-cognitive evidence	Behavioral engagement	Learning Time	
		Learning Concentration	
		Active Learning	
		Cooperative Learning	
		Classroom Participation	
	Cognitive engagement	Degree of Self-monitoring	
		Deep Cognitive Strategy Used	
	emotional engagement		Interest in Learning
			Confidence in Learning
			Value of Learning
Satisfaction			
Sense of Identity			

3.3. Integrating Quantitative and Qualitative Analyses to Uncover the Causal Links between the Two Types of Evidence

Research shows that there is a significant positive correlation between students' learning engagement and the level of learning ability, an increase in students' reading engagement can effectively improve reading ability. Teachers should synthesize cognitive and non-cognitive evidence, integrate the results of quantitative and qualitative analyses, and establish causal links between the two types of evidence, in order to characterize students' ability growth and generate growth paths more clearly. For cognitive performance evidence, a line graph can be used to show the characteristics of students' ability growth, and for non-cognitive evidence, a radar graph can be used to present the characteristics of students' ability performance. Comprehensively comparing the radar graph and the line graph can give a more complete picture and help further analyze the students' ability. Taking the novel units in the second volume of the compulsory textbook as an example, the line graph shows that Student 1's comprehension of the novel's theme grew rapidly, and the radar chart shows that the student excelled at cooperative learning and classroom participation dimensions.

Based on the results presented in **Table 2**, we not only find the strengths, weaknesses and growth trends of the student's novel reading ability in various dimensions, but also find a profound learning pattern that the increase in cooperative learning and classroom participation effectively improved the student's ability to comprehend the novel's themes.

4. Focusing and Using Evidence: Promoting Continuous Improvement of Teaching Based on Evidence

Focusing on and utilizing evidence can improve students' learning and teachers' teaching, and promote the positive cycle of "teaching-learning-assessment consistency".

4.1. Focusing on Evidence to Deepen Students' Participation in the Classroom

Presenting two kinds of evidence visually could discover the causal relationships and show the characteristics of development of students' vertical and horizontal abilities, which in turn allows students to pay more attention to the evidence of their own learning and promote the development of learning ability and meta-cognitive skills. Focusing on the evidence helps to efficiently carry out multiple evaluation methods combining self-evaluation, mutual evaluation and teacher evaluation in class, avoiding the adverse effects of subjective tendencies in teachers' teaching evaluation, and also helps students participate in activities in class with multiple perspectives, build a shared atmosphere, form an open mind, and promote classroom participation and cooperative learning to reach depth, and enhance the level of core literacy in high-quality reflections and dialogues to achieve the best Learning outcomes. At the same time, evidence-based evalua-

tion also facilitates teachers and parents to better assess the level of students' ability development and learning dilemmas, forming a multi-party educating synergy.

4.2. Utilizing Evidence to Enhance Teachers' Teaching Ability

Typical learning evidence reflects the common and individualized problems of language learning among students in a class, and reveals certain universal patterns of language learning. Teachers should "choose effective and even optimal teaching approaches based on reliable and authentic evidence, and help students to use proven learning methods (Li, 2021)", promote the evaluation of language teaching to go beyond the experience-based evaluation model of the past and enhance the scientific and relevant nature of teaching to make it possible to tailor teaching to students' needs. On this basis, teachers' "teaching cognitive ability", "teaching operational ability" and "teaching monitoring ability" should also be continuously improved (Shen & Wang, 2000), so as to ultimately promote the positive integration of "teaching-learning-assessment".

4.3. Promoting Delayed Evaluation with Evidence-Based Evaluation to Facilitate Teaching Improvement

The development of students' ability is characterized by a long period of time,, but the evaluation of classroom teaching is characterized by a typical immediacy. However, this contradiction can be effectively reconciled based on the concept of "evidence-based evaluation", through establishing rules of evidence, quantitatively analyzing the evidence of students' cognitive performance, developing qualitative evaluation tools, descriptively analyzing the evidence of students' non-cognitive, building a complete chain of learning evidence and promoting the orderly development of delayed evaluation. Promoting delayed evaluation with "evidence-based evaluation" can provide students with a more relaxed growth environment and a broader space for error tolerance, truly establish the main position of students in learning and help teaching improve in a smooth and healthy cycle.

5. Conclusion

To sum up, with the concept of "evidence-based assessment" as the guide, using delayed assessment as the assessment method, refining the typical evidence of students' learning in the learning process, and carrying out in-depth analysis based on quantitative and qualitative methods, the realization of "consistency in teaching and assessment" can be effectively promoted, thus enhancing students' core language literacy. This can effectively promote the realization of "consistency in teaching and assessment" and enhance students' core language literacy.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Feng, C. D. (2012). Introduction to the “Evidence-Centered” Educational Evaluation Design Model. *Shanghai Educational Research*, *No. 8*, 12-16.
- Li, X. (2021). Evidence-Based Education: A Practical Exploration in the United Kingdom. *Comparative Education Research*, *43*, 71-78.
- Ni, S. K., & Wu, X. C. (2011). Learning Engagement: Concepts, Measurement, and Related Variables. *Psychological Research*, *4*, 81-87.
- Ravindran, B., Gerun, B. R., & Dilbeck, T. K. (2005). Predicting Pre-Service Teachers’ Cognitive Engagement with Goals and Epistemological Beliefs. *Journal of Educational Research*, *98*, 222-232. <https://doi.org/10.3200/JOER.98.4.222-233>
- Shen, J. L., & Wang, K. R. (2000). On Teachers’ Teaching Ability. *Journal of Beijing Normal University (Humanities and Social Sciences Edition)*, *No. 1*, 64-71.
- Zhong, Q. C. (2012). The Challenge of Classroom Assessment. *Global Education Outlook*, *41*, 10-16.