Educational Technology Overview and Its Application in Essay Evaluation

Chenming Xue  
cxue34@gatech.edu

***Abstract—***In this paper, the educational technology status has been summarized. The representative and selected journals are reviewed, and the typical applications have been introduced. For a specific application: essay evaluation, research journals are explored and potential opportunities and resolutions are briefly proposed. This is the first exploration for this criteria and deeper information will be investigated later.

# 1 Introduction

The educational technology is "the study and ethical practice of facilitating learning and improving performance by creating, using, and managing appropriate technological processes and resources". (Robinson 2016) The learning objects are "any entity, digital or non-digital, that may be used for learning, education or training" (New York, 2002) covering content, pedagogical elements, and learning objects standards.

In practice this technology covers broad aspects including and not limiting to

1. synchronous and asynchronous (Mayadas, 1997), both fields have their advantages for example, synchronous provides online live education opportunity and asynchronous provides information related to design, decision and evaluation for each level.
2. Linear learning (Jean-Eric, 2013). Computer-based training (CBT) is conceptually similar to web-based training (WBT) which are delivered via Internet using a web browser and smart phones, tablets and laptops are usually used as devices. The assessments are easily scored and recorded by computer and online software, providing immediate end-user feedback and completion status. The lack of human interaction is claimed to be the drawback that needs to improve.
3. Collaborative learning (Stahl, 2006). Typically, it is about computer-supported collaborative learning (CSCL) concept that uses instructional methods designed to encourage or require students to work together on learning tasks. Buzz words about this are "e-learning 2.0" and "networked collaborative learning" (NCL), which rely on the advanced web technology offering the opportunities of sharing information between multiple groups of people. It is acclaimed to be "a breeding ground for creative and engaging educational endeavors." The collaborative learning is different from instruction normally from instructors/professors and increase studying efficiency.

Accompanying to the learning theoretical methods, hardware technologies are a critical part. In fact, only with the developments in the technologies such as internet, wireless, and computing capability etc., the new educational methods are realized. The hardware includes audio and video, computer, tablets and mobile devices, online collaborative and social learning, whiteboard, virtual classroom, and augmented reality.

Nowadays, a highly attractive technology for educational technology is artificial intelligence (AI) assistance learning (du Boulay,2015). The cutting edge technologies in this field include Artificial Intelligence in Education (AIED), Intelligent Tutoring Systems (ITS), Educational Data Mining (EDM) and Learning Analytics & Knowledge (LAK). It has been widely adopted in education from K-12 classrooms to college. Intelligent tutoring systems (ITSs) is one prominent AI-enhanced technology which can provide immediate and personalized feedbacks to students. In this bigdata ear the data driven cognitive and suggestive AI system will be highly effective in helping students learn.

To support teachers in a realistic, highly differentiated, self-paced classroom, remains an open research problem and one interesting and challenging field is automated essay scoring.

# 2 Auto mated Essay Scoring

## 2.1 Background

Unlike well-known objective automated multichoice scoring, the much challenging subjective automated Essay Scoring (AES) is defined as the computer technology that evaluates and scores the written essays, (Shermis, 2002) which has been developed to assist teachers in low-stakes classroom assessment and testing companies and states in large-scale high-stakes assessment. The benefits are time saving, cost/labor saving, reliability, fairness and generalizability in writing assessment. (Bereiter , 2003) In the previous exploration, there has been a number of studies to assess the accuracy and reliability of the AES systems with respect to writing assessment. There has been reported high agreement rates comparing to human raters. (Attali, 2004) However, the AES systems are in the early stage and have been criticized for such problems which need to improve: lack of human interaction, vulnerability to cheating, requiring a large set of sample text to train the system. The popular approaches will be reviewed and discussed, and new approaches will be raised in the following session.

## 2.1 Approaches discussion

**2.1.1 Project Essay Grader (PEG)**

PEG was firstly developed by Ellis Page in 1966 and planned to make the large-scale essay scoring process more practical and effective (Rudner, 2001). This method uses correlation to predict the intrinsic quality of the essays.

The terms trins and proxes were used for explaining the way PEG generates a score. Trins refer to the intrinsic variables (fluency, diction, grammar, punctuation, etc), and proxes denote to the approximation (correlation) of the intrinsic variables and refer to actual counts in an essay (e.g. establishing the correlation of fluency or trin with the amount of vocabulary or prox). The scoring system contains a training stage and a scoring stage. PEG is trained on a sample of essays in the former stage and in the latter proxy variables are determined for each essay and these variables are entered to the prediction equation. In the last stage, the score is assigned by computing beta weights (coefficients) from training stage. This method needs 100 ~ 400 sample essays for training.

This method has pros in predicting scores comparable to human raters and it can track wring errors. But this method is weak in evaluation in semantic aspect of essays and more focusing on the surface structures and failing to detect the content related features. (Kukich, 2000) Also, this method cannot provide instructional feedback to students.

## 2.1.2 Intelligent Essay Assessor (IEA)

Another AES system analyzes and scores an essay using a semantic text-analysis method called Latent Semantic Analysis (LSA) (Lemaire, 2001). This LSA method is defined as “a statistical model of word usage that permits comparisons of the semantic similarity between pieces of textual information”. The idea can be summarized as:

“meaning of word1 + meaning of word2 + … + meaning of wordk = meaning of passage”

This application of LSA including picking the most suitable text for students with different levels of background knowledge, automatic scoring of essay contents, and assisting students in summarizing texts successfully. The domain representative texts are required to train the LSA. In the algorithm of LSA, the text is represented as a matrix. Each row in the matrix represents a unique word, while each column represents context. Each cell involves the frequency of the word. Then each cell frequency is considered by a feature that denotes not only the importance of the word but also the degree to which the word type carries information in the domain discourse. The number of occurrences of each word in a text determines its semantic space, for example, 300 paragraphs and 2000 words provide a 300×2000 matrix. Reducing dimensions is critical and is induced by semantic similarities between words. If the number is too small, important information is lost, if the number is too big, limited dependencies will be drawn between vectors. LSA not only evaluates mainly the quality of the content of an essay, it also includes scoring and feedback on grammar, style and mechanics. Intelligent essay assessor (IEA) mainly focuses on the content related features rather than the forms. Detecting plagiarism is also an critical task as this is hard for humans. The architecture of IEA is shown in Figure 1.



**Figure 1.** The intelligent essay assessor architecture. Source: (Shermis, 2003).

The advantage of IEA is this system needs smaller numbers of pre-scored essays to train, like 100 pre-scored training essays.

### 2.1.3 E-rater and Criterion

Heading 1 should be set in all caps. It should have 11 points of space added before and 8.5 points of space added after.

### 2.1.4 IntelliMetric and MY Access

Besides *Heading 1,* which is set in caps, headings should always use sentence case (i.e., first word capitalized) rather than title case; after all, they are not titles. *Heading 2* should be set in bold roman (upright), and *Heading 3* should be set in bold italics. The use of headings beyond *Heading 3* is discouraged.

## 2.1.5 Bayesian Essay Test Scoring System (BETSY)

JDF uses the US Letter paper size (8.5″ × 11″). It has a top margin of 1″, and bottom and side margins of 1.5″. This yields a text block of 5.5″ × 8.5″, which is exactly ½ the size of the page, divided lengthwise.

The page number should be included in the bottom margin, 1″ from the bottom of the page – this creates symmetry with the top margin. No other elements should be placed in the margins.

# 2 PRESENTATIONAL ELEMENTS

You are encouraged to use presentational elements liberally, as long as they add to the clarity of your submissions. They often require less space and fewer accompanying words to explain a given concept, and do a far better job of it.

## 2.1 Figures

Figures should always be centered on the page, although they may also take up the entire width and height of the text block. Figures should always be referenced in the text, and they should include a descriptive caption. Figures may also be equations, diagrams, or other kinds of content.

If your figure includes a white background (e.g. an interface design or graph), it may aid legibility to add a ¼ point black border.

Figure captions should be centered beneath the corresponding figure. The label for the figure, e.g. “Figure 1,” should be bolded, and the entire caption should be 8.5 points with 14 points of line spacing. If need be, you may have one figure caption corresponding to multiple consecutive figures and use either locational descriptors (e.g. “top left,” “middle”) or labels (e.g. “A”, “B”) to map parts of the caption to parts of the figure. Make sure that caption falls on the same page as the corresponding figure or table; you may rearrange text to make this work.

## 2.2 Tables

You have freedom to format tables in the way that works best for your data. Generally, text should be left-aligned and numbers should be right-aligned or aligned at the decimal – you can do this using a [custom tab stop](https://practicaltypography.com/tabs-and-tab-stops.html). The default table style (shown below) reduces the text size to be equal to the caption text.

Table captions should be formatted the same way as figure captions, but they should be placed above the table. The popular mnemonic for this is: figures at the foot, tables at the top. Like figures, tables should not exceed the margins and should be centered on the page.

**Table 1.** Mathematical constants. Notice how the approximations align at the decimal.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Symbol** | **Approximation** | **Description** |
| Golden ratio | *φ* | 1.618 | Number such that the ratio of 1 to the number is equal to the ratio of its reciprocal to 1 |
| Euler’s number | *e* | 2.71828 | Exponential growth constant |
| Archimedes’ constant | *π* | 3.14 | The ratio between circumference and diameter of a circle |
| One hundred | A+ | 100.00 | The grade we hope you’ll all earn in this class |

## 2.3 Additional elements

There are additional elements you may want to include in your paper, such as in-line or block quotes, lists, and more. For other content types not covered here, you have reasonable flexibility in determining how it should be used in this format.

### 2.3.1 Quotes

If you would like to quote an outside source, you may do so with quotation marks followed by a citation. If a quote is fewer than three lines, you may write it in-line. It is acceptable to replace pronouns with their target in brackets for clarity. For example, “Heavy use of peer grading would compromise [the school’s] reputation” (Joyner, 2016). If a quote exceeds three lines, you should set it as its own paragraph with 0.5″ side margins, using the *Blockquote* style.

“Whether or not the grades generated by peers are reliably similar to grades generated by experts is only one factor worth considering, however. Student perception is also an important factor. A recent study indicated that reliance on peer grading is one of the top drivers of high MOOC dropout rates. This problem may be addressed by reintroducing some expert grading where possible.” (Joyner, 2016)

### 2.3.2 Lists

Bulleted and numbered lists are indented 0.5″ from the left margin, with the bullet or number hanging in the margin by 0.25″ (the default format).

# 3 PROCEDURAL ELEMENTS

## 3.1 In-line citations

Articles or sources to which you refer should be cited in-line with the authors’ names and the year of publication.[[1]](#footnote-1) The citation should be placed close in the text to the actual claim, not merely at the end of the paragraph. For example: students in the OMSCS program are older and more likely to be employed than students in the on-campus program (Joyner, 2017). In the event of multiple authors, list them. For example: research finds sentiment analysis of the text of OMSCS reviews corresponds to student-assigned ratings of the course (Newman *&* Joyner, 2018). You may also cite multiple studies together. For example: several studies have found students in the online version of an undergraduate CS1 class performed equally with students in a traditional version (Joyner, 2018a; Joyner, 2018b). If you would like to refer to an author in text, you may also do so by including the year (in parentheses) after the author’s name in the text. If a publication has more than 4 authors, you may list the first author followed by ‘et al.’ For example: Joyner et al. (2016) claim that a round of peer review prior to grading may improve graders’ efficiency and the quality of feedback given. This applies to parenthetical citations as well, e.g. (Joyner et al., 2016).

## 3.2 Reference lists

References should be placed at the end of the paper in a dedicated section. Reference lists should be numbered and organized alphabetically by first author’s last name. If multiple papers have the same author(s) and year, you may append a letter to the end of the year to allow differentiated in-line text (e.g. Joyner, 2018a and Joyner, 2018b in the section above). If multiple papers have the same author(s), list them in chronological order starting with the older paper. Only works that are cited in-line should be included in the reference list. The reference list does not count against the length requirements.

# 4 REFERENCES

1. **Joyner, D. A., Ashby, W., Irish, L., Lam, Y., Langston, J., Lupiani, I., Lustig, M., Pettoruto, P., Sheahen, D., Smiley, A., Bruckman, A., & Goel, A. (2016). Graders as Meta-Reviewers: Simultaneously Scaling and Im-proving Expert Evaluation for Large Online Classrooms. In Proceedings of the Third Annual ACM Conference on Learning at Scale. Edinburgh, Scotland.**
2. Attali, Y. (2004, April). Exploring the feedback and revision features of Criterion. Paper presented at the National Council on Measurement in Education (NCME), San Diego, CA.
3. Bereiter, C. (2003). Foreword. In Mark D. Shermis and Jill C. Burstein (Eds.), Automated essay scoring: a cross disciplinary approach (pp. vii–ix). Mahwah, NJ: Lawrence Erlbaum Associates.
4. Kukich, K. (2000, September/October). Beyond Automated Essay Scoring. In M. A. Hearst (Ed.), The debate on automated essay grading. IEEE Intelligent systems, 27–31. Retrieved November 12, 2004,
5. Robinson, Rhonda; Molenda, Michael; Rezabek, Landra. "Facilitating Learning" (PDF). Association for Educational Communications and Technology. Retrieved 18 March 2016.
6. Lemaire, B. & Dessus, P. (2001). A system to assess the semantic content of student essays. Educational Computing Research, 24(3), 305–306.
7. Mayadas, F (March 1997), "Asynchronous learning networks: a Sloan Foundation perspective", Journal of Asynchronous Learning Networks,
8. "Collaborative asynchronous online learning". US Patent Office. 10 March 2014. Retrieved 23 March 2019. This article incorporates text from this source, which is in the public domain.
9. Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), Cambridge handbook of the learning sciences (pp. 409-426). Cambridge, UK: Cambridge University Press.
10. Learning Technology Standards Committee (2002), Draft Standard for Learning Object Metadata. IEEE Standard 1484.12.1 (PDF), New York: Institute of Electrical and Electronics Engineers, retrieved 2008-04-29.
11. du Boulay, Benedict (2015-08-06). "Recent Meta-reviews and Meta–analyses of AIED Systems". International Journal of Artificial Intelligence in Education. 26 (1): 536–537. doi:10.1007/s40593-015-0060-1. ISSN 1560-4292
12. Rudner, L. & Gagne, P. (2001). An overview of three approaches to scoring written essays by computer (ERIC Digest number ED 458 290).
13. Shermis, M. & Barrera, F. (2002). Exit assessments: Evaluating writing ability through Automated Essay Scoring (ERIC document reproduction service no ED 464 950).
14. Shermis, M. D., Raymat, M. V., & Barrera, F. (2003). Assessing writing through the curriculum with Automated Essay Scoring (ERIC document reproduction service no ED 477 929).

# 

# 5 APPENDICES

You may optionally move certain information to appendices at the end of your paper, after the reference list. If you have multiple appendices, you should create a section with a *Heading 1* of “Appendices.” Each appendix should begin with a descriptive *Heading 2*; appendices can thus be referenced in the body text using their heading number and description, e.g. “Appendix 5.1: Survey responses.” If you have only one appendix, you can label it with the word “Appendix” followed by a descriptive title, e.g., “Appendix: Survey responses.”

These appendices do not count against the page limit, but they should not contain any information required to answer the question in full. The body text should be sufficient to answer the question, and the appendices should be included only for you to reference or to give additional context. If you decide to move content to an appendix, be sure to summarize the content and note it in relevant place in the body text, e.g., “The raw data can be viewed in *Appendix 5.1: Survey responses.*”

1. In-line citations are preferred over footnotes, and we favor APA citation format for both in-line citations and reference lists. Refer to the [Purdue Online Writing Lab](https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_guide/in_text_citations_the_basics.html), or follow the above examples. Footnotes should use 8.5 point text with 1.26 line spacing. [↑](#footnote-ref-1)