AI AND WRITING: FROM 1950 TO TODAY

The Turing Test (the "imitation game") is a central concept in AI and used as a benchmark to assess the intelligence of machines.

Can machines think?

1950

1965

Can machines interact?

Joseph Weizenbaum finishes ELIZA, a natural language processing environment. Its most famous mode was called DOCTOR, which responded to user questions much like a psychotherapist.

1966

Can machines grade?

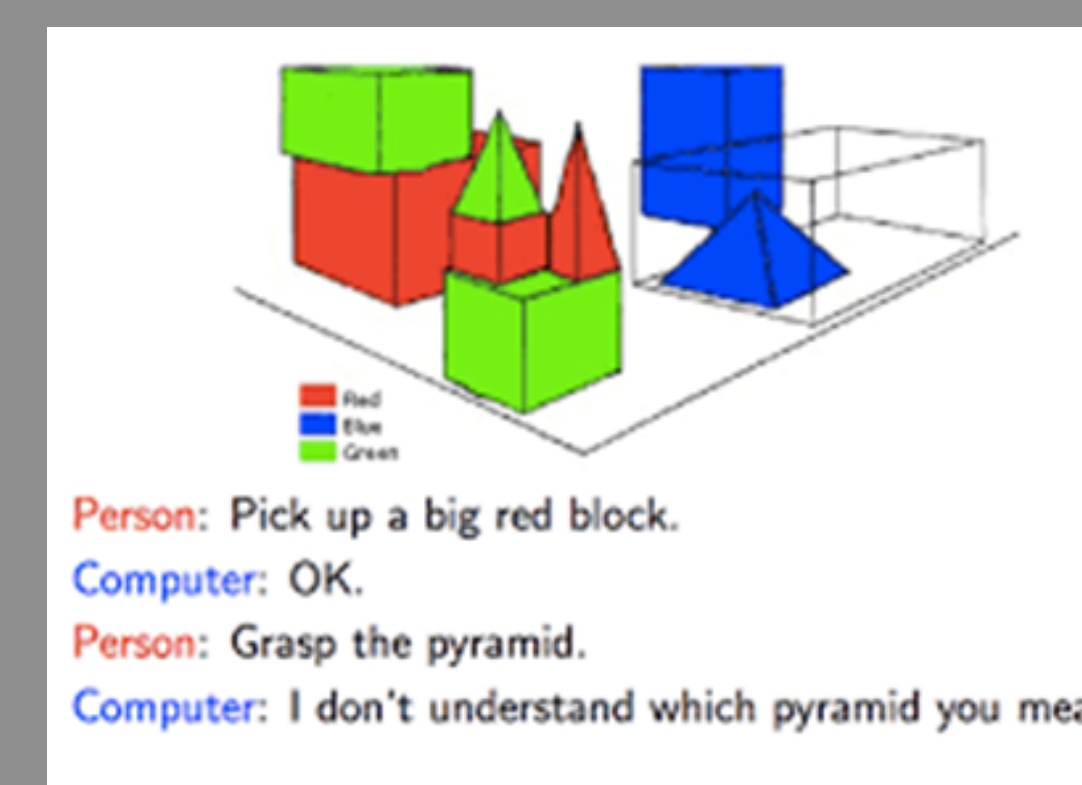The Imminence of . . .
Grading Essays by Computer

Breakthrough? Or buncombe and ballyhoo?
You should know, after reading this careful description of efforts at the University of Connecticut to rescue the conscientious English teacher from his backbreaking burden. It is authored by the researcher whose very first computer strategy for essay grading yielded marks indistinguishable from those of experts. Mr. Page, himself a refugee from English teaching, answers questions that will occur to the skeptical educator.

By ELLIS B. PAGE

Nonsense!
Hogwash!
Ballyhoo!
Impossible!

1970

Can machines understand humans?

SHRDLU, developed by Terry Winograd, can interact with people in English and move objects in a virtual block world. This laid the groundwork for Natural Language Processing (NLP).
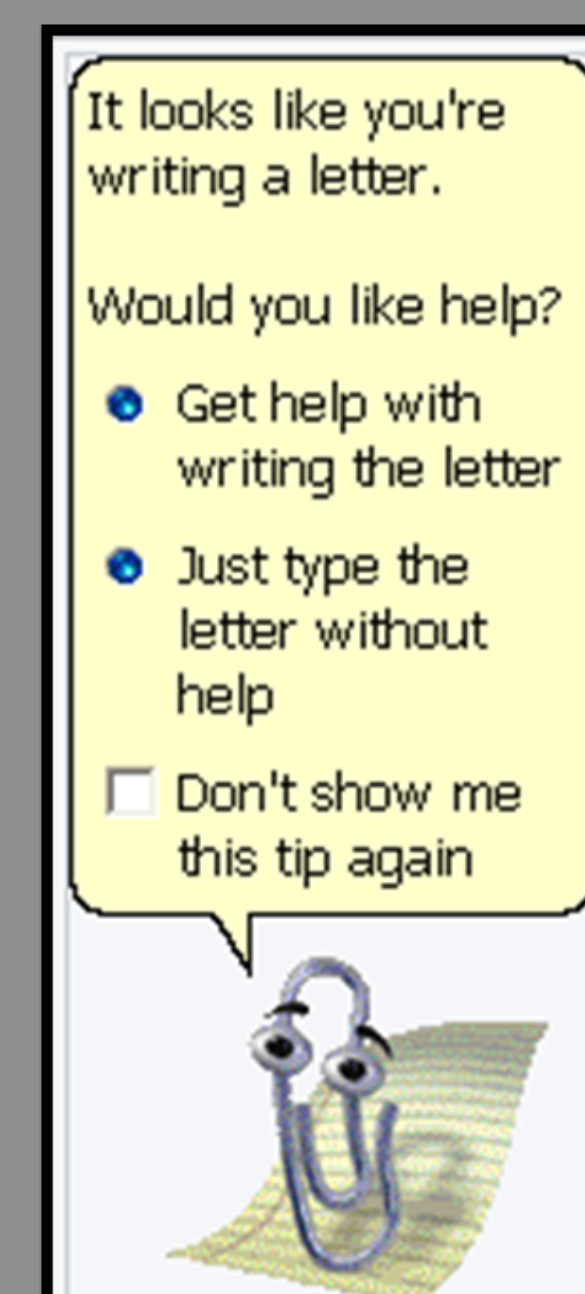
Person: Pick up a big red block.
Computer: OK.
Person: Grasp the pyramid.
Computer: I don't understand which pyramid you mean

PEG writing®
*ets
e-rater®
Criterion®
AES
AEE
AWE

Can machines score (reliably)?

1980-1990

1990-2000

Can machines help us write?

It looks like you're writing a letter.
Would you like help?
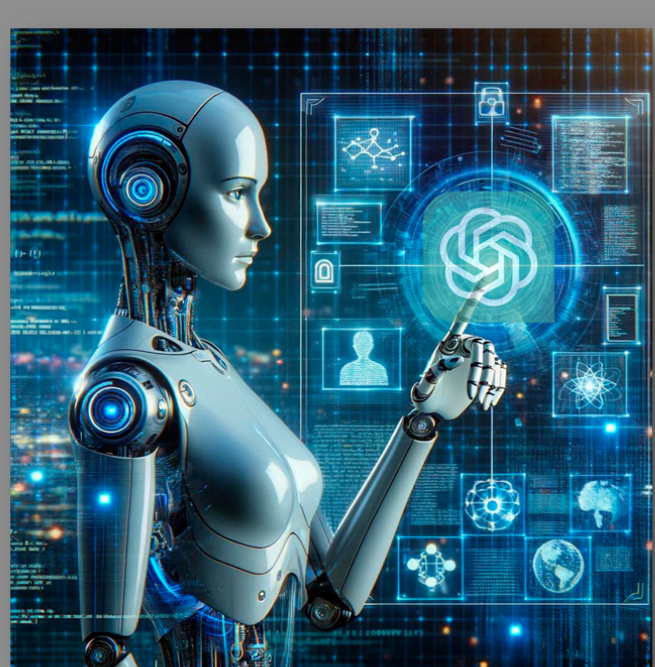• Get help with writing the letter
• Just type the letter without help
☐ Don't show me this tip again

The probability of this essay coming from group 1 is 0.9326
The probability of this essay coming from group 2 is 0.9674
The probability of this essay coming from group 3 is less than .01
The probability of this essay coming from group 4 is less than .01
The probability of this essay coming from group 5 is less than .01
Therefore, I am very certain that the essay belongs in group 2

[most likely]

The probability of this essay coming from group 1 is 0.5989
The probability of this essay coming from group 2 is 0.3993
The probability of this essay coming from group 3 is less than .01
The probability of this essay coming from group 4 is less than .01
The probability of this essay coming from group 5 is less than .01
Therefore, I think the essay belongs in group 1

Generative AI (GAI) and deep learning networks (transformers) can generate text and images that mimic human writing. These models now learn and create.
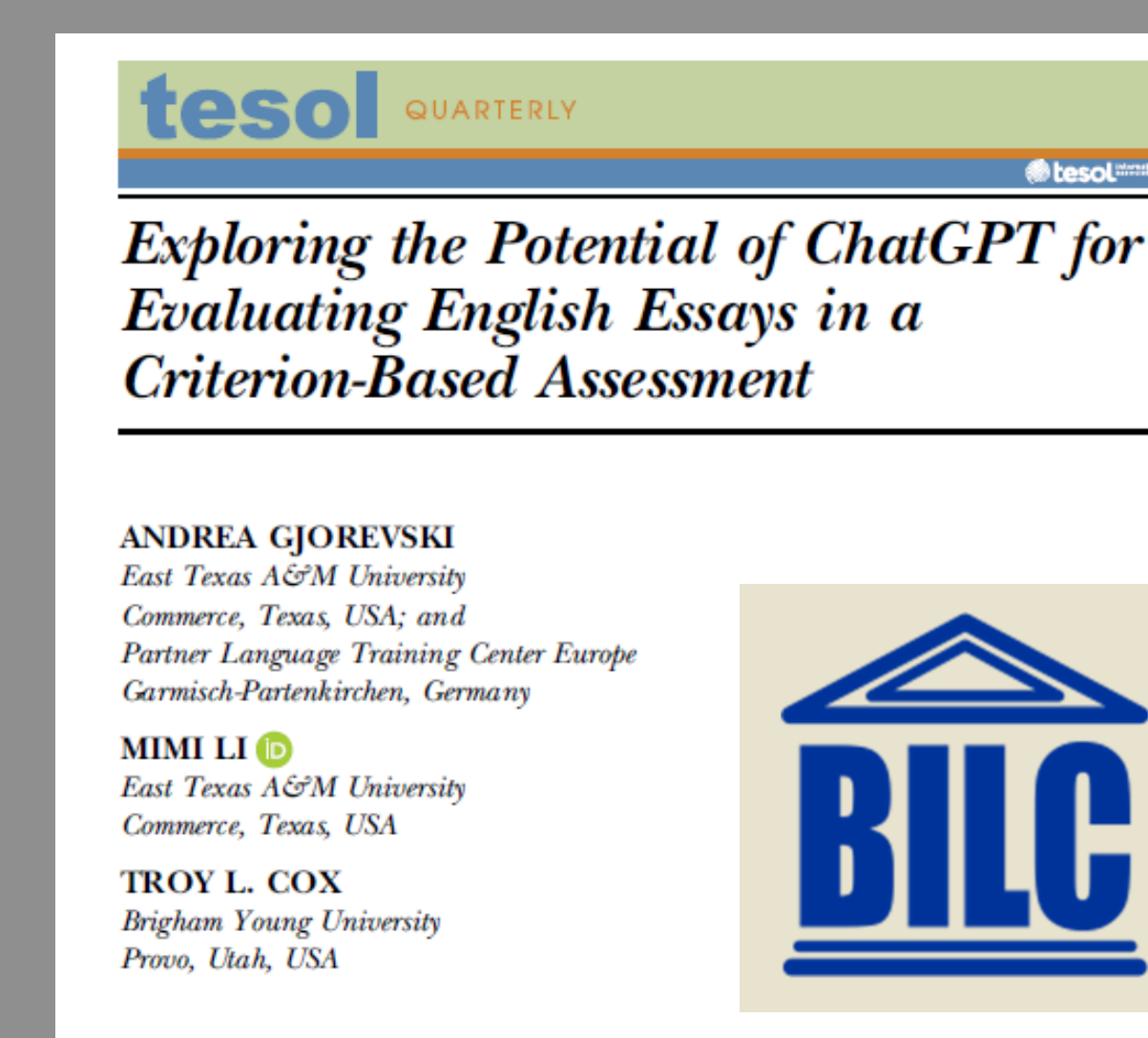
Can machines write like humans?

2010-TODAY

Can machines score writing like humans?

2025

tesol QUARTERLY
Exploring the Potential of ChatGPT for Evaluating English Essays in a Criterion-Based Assessment

ANDREA GJOREVSKI
East Texas A&M University Commerce, Texas, USA; and Partner Language Training Center Europe Garmisch-Partenkirchen, Germany

MIMI LI
East Texas A&M University Commerce, Texas, USA

TROY L. COX
Brigham Young University Provo, Utah, USA

BILC

BILC testers and ChatGPT rated 100 essays and provided justifications for their ratings. Results showed exact agreement 50%; adjacent agreement 91%. Conclusion: substantial training is required, but ChatGPT showed some understanding of the scale. More research is needed, but there is potential!

Sources:
https://www.ibm.com/think/topics/history-of-artificial-intelligence
TESOL Quarterly, 3 August 2025, https://doi.org/10.1002/tesq.70011