

THE CONVERSATION

4 May 2015, 1.07pm AEST

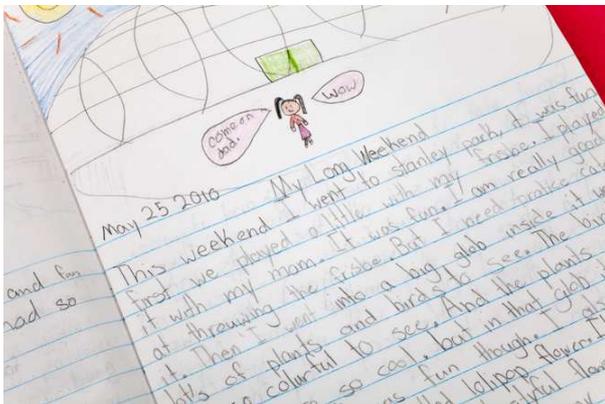
Who needs teachers when computers can mark exams?

AUTHOR



Stewart Riddle

Senior Lecturer at
University of Southern
Queensland



Will a computer algorithm understand the creativity, flair, irony, wit and unconventional approaches used in kids' writing? from www.shutterstock.com.au

Early May would be incomplete without some NAPLAN controversy. This year's comes from the announcement last week that the national exam sat by students across the country in Years 3, 5, 7 and 9 is to be marked by computers in 2017.

Part of the argument for moving to online marking is that it will decrease turnaround time from months to just weeks. While this is uncontroversial for multiple-choice-style tests, which have a correct answer, it is much more problematic when applied to creative writing.

Can computers mark creative writing?

The NAPLAN written task is usually a narrative or persuasive task and is an extended piece of prose. The marking criteria include audience, text structure, cohesion, vocabulary, paragraphing, sentence structure, punctuation and spelling.

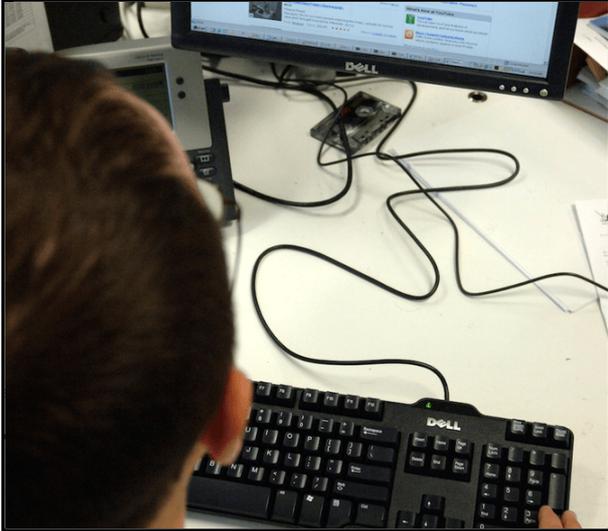
When writing persuasive texts, the guide explains that:

students are required to write their opinion and to draw on personal knowledge and experience when responding to test topics.

The guide also explains that for narrative texts, there should be a:

growing understanding that the middle of the story needs to involve a problem or complication that introduces conflict, danger or tension that must be resolved. It is this uncertainty that draws the reader in and builds suspense.

The question is whether computers can appropriately mark students' creative writing with this level of sophistication.



Does an algorithm marking system teach our kids what matters in education? AAP

[Click to enlarge](#)

According to the Australian Curriculum, Assessment and Reporting Authority (ACARA), they can.

The approach being taken is one that uses supervised machine learning, where sample tests marked by humans are fed into an algorithm that learns how to recognise quality responses by reverse-engineering scoring decisions. Trials conducted by ACARA have demonstrated that:

artificial intelligence solutions perform as well, or even better, than the teachers involved.

One argument is that computer marking has less variability than human markers, although these claims to marker reliability are contested.

For example, what would happen if a student were to submit a nonsense piece that happened to meet the expectations of the algorithm?

Automated marking is not a new thing. It has been particularly visible since the rise of MOOCs and the search for a cheap alternative for marking student papers.

The research literature provides a mixed picture of potential benefits and pitfalls, yet there has been vocal opposition to computer marking from academics and educationalists.

The rise of algorithms can be seen in many places, including chess-playing computers, self-driving cars, metadata analysis to predict behaviour, online advertising, speech-recognition software and auto-completing search engines. It seems only logical that algorithms would enter our classrooms.

What actually matters in education?

One thing that strikes me as ironic is that we would be using computers, which can't actually read or write, to test the reading and writing of our students. Is the next step to replace our teachers with robot instructors who can provide standardised, objective and completely emotionless feedback in the classroom?

How can a computer assess creativity and flair? How would it recognise irony, wit and humour? What about writers who use unconventional approaches for effect?

While algorithms can easily process literal meaning, what happens with inferential meaning or drawing on rich contexts, background knowledge, prior learning, cultural and social discourses? These are all part of the complex tapestry of human meaning-making in reading and writing.

As one example, the NAPLAN marking guide refers to the use of classical rhetorical discourse in persuasive writing, including:

Pathos - appeal to emotion

Ethos - appeal to values

Logos - appeal to reason.

I have not yet come across a computer except in science-fiction films that has emotions or values that could be appealed to in any persuasive sense.

There are serious concerns that computer marking of the NAPLAN writing task will have unintended effects on **teaching and learning**, including online reading and writing **strategies** different to those of traditional print-based comprehension and composition.

A further concern is that computer marking will have a reductive effect on student writing, with "teaching to the test" becoming more of a problem than it already is.

Maybe it isn't that far-fetched to imagine computers marking assignments and robots teaching classrooms. After all, there are **predictions** that we will reach the singularity, the point at which artificial intelligence overtakes humans, in 2029.

Wouldn't ACARA be better off putting the money into something that has an impact on the quality of learning of students in Australian schools rather than conducting this particular experiment? To be focusing on test scoring that is faster and cheaper seems to be at odds with what actually matters in education.

Until we reach the singularity, perhaps we should focus on improving equity and access for students who are most disadvantaged in our education system, and leave the robots out of it.