

# The Effectiveness of Free Software for Detecting AI-Generated Writing

Gregory Price<sup>1</sup>, Marc D. Sakellarios<sup>2</sup>

<sup>1</sup>Faculty of Letters, Chuo University, Japan

Email: [dprice001@g.chuo-u.ac.jp](mailto:dprice001@g.chuo-u.ac.jp)

<sup>2</sup>Department of Science and Engineering, Tokyo University of Science, Japan

Email: [msakell@rs.tus.ac.jp](mailto:msakell@rs.tus.ac.jp)

Received: 03 Oct 2023, Received in revised form: 08 Nov 2023, Accepted: 15 Nov 2023, Available online: 23 Nov 2023

## Abstract

*As artificial intelligence (AI) technology advances, educators are faced with the challenge of detecting AI-generated writing in student assignments. This study explores the effectiveness of freely available AI detection software in identifying AI-generated content and its implications for teacher evaluation of student work. The research involved Japanese university students who used AI for English writing assignments. We compared our manual analysis with results from five different free AI writing detectors. The findings reveal the limitations of free AI detection tools and the need for cautious use in educational settings.*

**Keywords—** *AI-generated writing, AI detection software, education technology, free AI tools, student evaluations.*

## I. INTRODUCTION

As the technology of artificial intelligence develops rapidly, it could perhaps be suggested that humans are at the start of a new age in computer-assisted thinking and writing. English language educators tasked with appraising student writing may be particularly interested in how to deal with this emerging technology in the context of their classrooms. Teachers wishing to use AI detection tools may not have equal access to paid software or registration fees. Therefore, the research question that we are exploring in this paper concerns software that is available to all without the exclusivity of a paywall. How effective is free AI detection software for the recognition of AI-generated writing and is it conclusive enough to support reliable teacher evaluations of student work?

## II. LITERATURE REVIEW

The question regarding whether machines can think can perhaps be traced back to Alan Turing's seminal

1950 work titled "Computing Machinery and Intelligence." Instead of asking whether machines and computers can think, Turing decided to posit a more pragmatic question as to whether computer speech could be indistinguishable from that of humans (Zemčík, 2019). Turing refers to this concept as "The Imitation Game" (Turing, 2012). Turing (2012) predicted that future machines will play "the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning" (p. 449). In the case of AI-generated writing, the interrogator cannot engage in back-and-forth dialogue, but must rather analyze the human Imitation Game based on a final product.

In November 2022, the American company OpenAI developed ChatGPT, which is considered to be the most advanced AI bot to date (Alafnan et al., 2023). In the first two months of ChatGPT's release, it gained over 100 million active users to break the record for the most rapid expansion of any application in history

(Buriak et al., 2023; Wu et al. 2023). In March 2023, GPT-4 was released by OpenAI to further enhance the useability of the software (Wu et al. 2023). AlAfnan et al. (2023) describe the advancement of this new AI technology as a nightmare for teachers faced with grading computer-generated writing being passed as human work. It can be assumed that teachers presented with writing will analyze its humanity either through their own intuition, or AI detection programs. We will briefly discuss these two methods of AI detection from here on.

Teacher intuition when appraising the humanity of writing may be ineffectual. Yan et al. (2023), citing Clark et al. (2021) and Ippolito et al. (2019) relay that, based on experiments, “humans rarely perform better than random guessing (60% - 65%) when asked to identify texts generated by modern AIs” (p. 126). Additionally, human appraisers of writing provide arbitrary rationales for determining whether writing is human or AI-generated. Human appraisers will generally evaluate writing based on whether it rambles or feels natural (Yan et al., 2023). Dugan et al. (2020) suggest that humans can be trained to better detect AI writing, but their accuracy after training is still very low.

Software aimed at detecting AI offers another avenue for teachers wishing to ascertain the humanity of a piece of writing. Walters (2023) describes a process often used by AI detection programs. The writing is broken down into words, or tokens. The programs then predict how one token will be typically followed by another. Following that, the programs rate the perplexity and idiosyncratic nature of the writing. The working concept is that humans tend to write with more perplexity and idiosyncratic feel than a machine (Walters, 2023).

Walters (2023), offering meta-analysis of data generated by 13 studies aimed at assessing AI detection software (Aremu, 2023; Cingillioglu, 2023; Desaire et al., 2023; Gao et al., 2023; Guo et al., 2023; Khalil & Er, 2023; Krishna et al., 2023; Liang et al., 2023; Pegoraro et al., 2023; Perkins et al., 2023; Wang et al., 2023; Weber-Wulff et al., 2023; Yan et al., 2023) found that “GPT-2/RoBERTa, TurnItIn, and ZeroGPT are the most consistently accurate detectors” (p. 5). Walters (2023), in his own study analyzing the effectiveness of 16 publicly available AI detection software found that CopyLeaks, TurnItIn, and Originality.ai have a high accuracy in detecting AI-generated writing. The 13 other programs were accurate at distinguishing human writing from GPT 3.5,

but not the newer GPT4.0 (Walters, 2023). Walters (2023) also found that pay software was slightly more accurate than the free counterparts.

Finally, adding to this issue is the advent of AI grammar-checking software such as Grammarly, Ludwig, Ginger, etc. Fitria (2021) explains that grammar-checking software like Grammarly is “powered by an interconnected system that incorporates rules, developments, and techniques of artificial intelligence such as machine learning, deep learning, and processing of natural languages” (p. 67). This calls into question whether sentences devised by AI-generated grammar-checkers constitute original writing. Grammar-checking software such as Grammarly requires human feedback in the form of clicking “ignore suggestion” from the human subject (Fitria, 2021). If the student is suggested a more eloquent sentence, do teachers expect the student to click “ignore” to produce truly original sentences? If the teacher is grading the student based on their writing ability, perhaps grammar-checking software is giving the teacher a skewed sample of the student’s true writing.

### III. METHOD

The findings for this paper come from a semi-structured survey and post-survey interviews related to a month-long study. The participants (n=3) were three Japanese-national English-as-a-foreign-language (EFL) students at a liberal arts university in Tokyo. The class was an elective and the students ranged from one first-year student to two fourth-year students. The class was not mandatory; the fact that students voluntarily registered for the class indicates their motivation to develop their English skills.

The participants were aware of the study from the first day of the semester. We explained the purpose of the research was to examine aspects of AI-generated writing as used for assignments by university students. The students were given four topics for writing and discussion spread out over several weeks. The teacher gave the students the following directives: 1.) Use AI to generate writing for at least one of the assignments, 2.) Write at least one of the assignments completely without the use of AI, and 3.) Do not divulge which assignment was AI-generated until the fifth week when they were asked to answer survey questions and take part in oral interviews related to the study.

The topics used were a mixture of global and national current affairs. These included population (growth and

depopulation), economics (inflation vs. deflation), pros and cons of genetically modified foods, and potential revisions to the Japanese Constitution. The lessons were designed in a flipped class structure where the students were informed of the topic and given a week to prepare essays for use in discussions in the following class session.

For the purposes of detecting AI writing, we used five free AI writing detectors that were readily found online. The free software tools used were Content at Scale AI Detector, GPTZero, Sapling AI Detector, Writer AI Content Detector, and ZeroGPT.

The detection process did not solely rely on the software. We analyzed each of the assignments and looked for clues as to whether AI writing was used. As the participants were non-native language users, some identifying markers could suggest that the writing was human-generated. These would include grammatical inconsistencies such as misuse of articles, or unnatural vocabulary choices. All students had submitted other unrelated assignments earlier in the semester and the two fourth-year students had previously had the same teacher for academic writing courses. Therefore, we had a baseline of material from which to compare writing styles.

With each submitted essay, we examined them first using manual analysis and then by selecting sections of the writing and entering them into the five free AI writing detectors. Occasionally, we would run tests on entire essays, but these never showed definitive results for AI-generated writing, as all the students included some human-created writing in parts of the essays, such as in the introductions and conclusions.

After the four sessions, the students were given a survey (see Appendix A) with questions eliciting them to explain their use of AI-generated writing for these assignments. Questions 2-4 were especially relevant to the study reported in this paper. Question 2 asked which AI software they chose to use. Question 3 asked which essay topics they chose to incorporate AI-generated writing for their assignments. Question 4 expanded on those by asking how they used the software. The responses to these three questions are useful to understand the results of our attempts at detecting the use of AI-generated writing for these assignments.

#### IV. RESULTS

The participants used AI-generated writing for at least one of their assignments. In the semi-structured interview, the participants explained which of the assignments they used AI-generated writing to assist with their work and to what degree. In this section, we will show a breakdown of the results from the five AI detection tools for one AI-generated essay from each of the participants.

The results listed here come from a process in which the teacher first examined the writing, and then copied the body of the essay into the five AI detection tools. The introduction and concluding sections of the essays were not used because we assumed there was a possibility that the students included human-written content in those sections to personalize the overall tone of the essay. The following results indicate the teacher's notes as written before knowing the method of writing (AI or human), as well as the results from the AI detection tools. Though the tools we used are listed here, the purpose of this paper is not to support or denigrate any particular software but rather to give an overall perception of the effectiveness of free software for detecting AI-generated writing.

Participant 1 used AI-generated writing for the assignment about population problems. In the survey and interview, they explained their process using ChatGPT. The essay answer that they submitted had very few adjustments to what was generated by AI, even though they said that they disagreed with some of the opinions expressed in the AI-generated essay.

The teacher's first impression of the essay was that it likely was written by AI. The teacher wrote, "I think it is written by AI, based on the tags which begin each new paragraph." These numbered tags are indicators that Chat GPT regularly uses to highlight each main point (see Appendix B). The results from the free online detection tools were as follows. Content at Scale AI Detector said there was a 98% probability that the writing was AI-generated. GPTZero listed the possibility as 56%. Sapling AI Detector listed it as 99.7%. Writer AI Content Detector alternately determined that only 21% of the material was written by AI. ZeroGPT stated that 0% of the writing was AI-written.

Participant 2 used AI-generated writing for the assignment about changes to the Japanese Constitution. In the interview, they stated that they used AI for about 50% of the assignment. They explained that they used their own opinion, but included explanations generated

by AI. Their process involved using ChatGPT as well as Grammarly.

The teacher's notes indicate that they detected some AI writing in use. One main indication the teacher noted was that there was a variation in capitalization of the word "constitution" when discussing the Japanese Constitution and constitutions in general. This discrepancy was the main clue that it was not fully generated by the same writer. The results from the free online detection tools were as follows. Content at Scale AI Detector listed the probability of AI-generated writing at 0%. GPTZero determined the possibility of AI generation to be 2%. Sapling AI Detector declared it to be 0%. Writer AI Content Detector stated that 0% of the material was written by AI. ZeroGPT also claimed 0%.

Participant 3 used AI-generated writing for the assignment about economics (inflation vs. deflation). Participant 3 explained in the interview that they developed the essay in sections by asking specific questions to ChatGPT in Japanese and then translating the responses into English on their own. In the interview, the student mentioned that part of their goal was to "deceive" the teacher as to their use of AI. Translating the AI-generated responses from Japanese to English on their own was a part of that strategy.

The teacher's notes indicate a leaning towards the possibility of help with AI, but the personal voice was strong enough that it was unlikely to be completely written by AI. The results from the free AI detection tools show that Participant 3's approach was effective. Content at Scale AI Detector listed the probability of AI-generated writing at 0%. GPTZero said the possibility was 19%, stating that it was most likely created as a combination of human and AI writing. Sapling AI Detector asserted that 0% was AI-generated content. Writer AI Content Detector declared that 0% of the material was written by AI. ZeroGPT also stated that 0% of the writing was AI-written.

## V. DISCUSSION

The results shown above indicate that the free AI-generated writing detection tools were not universally accurate. In the case of Participant 1, which represents the most fully AI-generated essay, the results were the most accurate. In the case of Participants 2 and 3 where a wider degree of variations was incorporated to both mask the use of AI as well as more accurately represent their own opinions on the topics, the

detection tools were not accurate. In all cases, the teacher's assessments were more accurate.

Until the end of the process, we did not know for sure which of the essay assignments the students were generating with the aid of AI. The participants wrote four assignments. We analyzed them for AI-generated writing first by our own reading, then by inputting selections (or sometimes full essays) through all five free AI detection tools. It would perhaps be outside the scope of this discussion to explain each of the seventy-two results, so we will focus on the assignment that was most fully generated by AI.

Participant 1 used AI for only the first question about population problems (see Appendix B). They used ChatGPT. The question posed by the teacher was "Do you think that there is a population problem? If 'yes', how can we fix it? If 'no', why do you think it is not a problem?" They reported in the semi-structured interview that they pasted this question directly into ChatGPT. The results were submitted with little adjustment.

The teacher first analyzed Participant 1's essay without tools. They noted that there was no referenced information given in the assignment. The essay was neatly arranged, starting with an introductory paragraph agreeing that there was a population problem (in Japan) and then listing a series of four methods to deal with the situation. The solutions were mapped out with numbered headings for each paragraph. This detail indicated that AI was likely used for generating this work, as similar writing constructions have been noticed in the researchers' own experiences using ChatGPT in preparation for this research.

The language and grammar used in the essay were quite natural for a non-native speaker. It is difficult to use this as a definitive cue as it could be that the student put extra effort into the work. However, certain indicators were shown here, such as a prevalent use of colons and a complete lack of any grammatical errors (such as misuse of articles) or unnatural vocabulary. A comparison with previous unrelated writing by the same student reinforced the impression that the essay was written either partially or completely by AI.

Of the five detection tools used for this study, a majority indicated that the assignment was likely to have been AI-generated. However, as this one particular example employed complete use of AI in its creation, the data is not strong enough to indicate that free AI detection software can be fully relied upon for

screening student assignments. The results for Content at Scale AI Detector were a 98% probability of AI generation. GPTZero showed a 56% probability. Sapling AI Detector was strongest with the claim of 99.7% probability for the use of AI. Writer AI Content Detector was 21%, and ZeroGPT stated that it was 0% written by AI. So, combined we get a claim of 54.94% probability that the assignment was generated by AI. The margin of error makes this software inconclusive in the evaluation of a student assignment.

There were very few cases of false positives. Mostly the AI detection results suggested human-created content. There was no case in which the teacher thought the work was AI when it was written completely by a human. One of the few examples of a false positive would be from an essay written by Participant 1 (see Appendix C). For this essay the teacher's notes state, "This has some odd wording choices and issues with grammar. That and the inclusion of a non-standard reference list give me the impression it was written by a human." Most of the AI detectors agreed that it appeared to be human. However, the Sapling AI Detector gave a 2% probability for the use of AI. Writer AI Content Detector was stronger with 15%. These are low but support the idea that there is a margin of error. This is acknowledged in some form on each of the free AI detection web pages. For example, in the instructions section of the webpage for the Sapling AI detector it is stated, "No current AI content detector (including Sapling's) should be used as a standalone check to determine whether text is AI-generated or written by a human. False positives and false negatives will occur" (Sapling, 2023, para 2)

### 5.1 Further research

The authors of this paper feel that there are more issues with AI-detection software that need addressing in future research. As shown above, AI-detection software uses a percentile scale to analyze "the level of human-like or AI-like quality of the essays" (Aremu, 2023, p. 3). If, for example, a student turns in a paper that is determined to be 95% likely to be AI-generated, should the teacher give the student the benefit of the doubt along the 5% margin of error? This situation is much different from direct plagiarism where a source can be located, and plagiarism can be determined with 100% accuracy.

Further compounding the problem of assessing whether a particular student has genuinely produced a piece of original writing is the increase in modern methods of teaching such as distance learning and

flipped classrooms. Distance makes proctoring tests more complicated. Khalil & Er (2023) suggest that the increase in distance education spanning the past 2 decades has caused institutions to "forfeit control over their digital educational infrastructure" (p. 3). This may, in turn, damage or call into question the institution's reputation (Khalil & Er, 2023). How should expanding technology be addressed by institutions that must choose to embrace major changes or navigate the definitions of divergent tools for use by students in academic environments?

Finally, what are the students' perceptions of using these tools in an academic environment? One main purpose of education is to prepare the students for the future. Do students want to embrace technology that will likely be prevalent in the society they will engage with? Do they accept current norms of academic practice as relevant? These are questions that require further research.

## VI. CONCLUSION

This report has sought to gain a clearer understanding of whether free AI-generated writing detection tools are reliable for teachers to use when considering student assignments. The participants in this research were Japanese university students writing in EFL. Other than the added clues this provides the teacher for human analysis, the results of this study can be useful for any teacher debating the choice to use free software tools for recognizing AI-generated writing for student assignments.

This study did not incorporate software that involves any sort of payment. There are several such detection tools, yet not all educators have access to them. Therefore, for this research, we have chosen to disregard information relating to such tools. It is our opinion that any tools, whether free or paid software, should only be used lightly to provide support for or against suspicion of student dishonesty. By no means should any tool be used for a definitive verdict in such cases.

Originally, we started with six tools for AI detection. We dropped consideration of the results from one of the tools, as during this study the software was discontinued by its creator. AI Writing Check.org (2023) provides the reason for discontinuing the free AI detection tool on their website.

The new versions of Generative AI tools are too sophisticated for detection by AI... and

each of these tools is being upgraded weekly. As these tools make their AI more complex, the AI text output becomes more varied, and it becomes more difficult for algorithms to detect whether a piece of writing was generated by AI. (August 2023 Update, para. 1)

Artificial Intelligence is a burgeoning technology, and there will be developments on both sides in the push-and-pull arena of AI detection. There are many elements of AI-generated writing in education that will presumably be researched and reported on. Some angles for future research would be explorations into techniques to deal with the difficulties associated with attempts to detect AI-generated work in students' writing as well as a study of the evolution of societal and academic opinions regarding the uses of such technology. We plan to explore more research findings from this study to gain a deeper understanding of the current perceptions of students regarding the use of AI in their schoolwork. The incorporation of AI-generated writing in academia is likely to be a fascinating field of study throughout the coming years.

## REFERENCES

- [1] AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). Chatgpt as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, 3(2), 60-68. <https://doi.org/10.37965/jait.2023.0184>
- [2] AI Writing Check. (2023). *August 2023 Update - AI Writing Check is no longer available* [Web page]. <https://aiwritingcheck.org/>
- [3] Aremu, T. (2023). Unlocking Pandora's Box: Unveiling the Elusive Realm of AI Text Detection. <http://dx.doi.org/10.2139/ssrn.4470719>
- [4] Buriak, J. M., Akinwande, D., Artzi, N., Brinker, C. J., Burrows, C., Chan, W. C., ... & Ye, J. (2023). Best Practices for Using AI When Writing Scientific Manuscripts: Caution, Care, and Consideration: Creative Science Depends on It. *ACS nano*, 17(5), 4091-4093. <https://doi.org/10.1021/acsnano.3c01544>
- [5] Cingillioglu, I. (2023). Detecting AI-generated essays: the ChatGPT challenge. *The International Journal of Information and Learning Technology*, 40(3), 259-268. <https://doi.org/10.1108/IJILT-03-2023-0043>
- [6] Clark, E., August, T., Serrano, S., Haduong, N., Gururangan, S., & Smith, N. A. (2021). All that's human is not gold: Evaluating human evaluation of generated text. *arXiv preprint arXiv:2107.00061*. <https://doi.org/10.48550/arXiv.2107.00061>
- [7] Content at Scale (2023). AI Detector. [Web-based software]. <https://contentatscale.ai/ai-content-detector/>
- [8] Desaire, H., Chua, A. E., Isom, M., Jarosova, R., & Hua, D. (2023). Distinguishing academic science writing from humans or ChatGPT with over 99% accuracy using off-the-shelf machine learning tools. *Cell Reports Physical Science*. <https://doi.org/10.1016/j.xcrp.2023.101426>
- [9] Dugan, L., Ippolito, D., Kirubarajan, A., & Callison-Burch, C. (2020). RoFT: A tool for evaluating human detection of machine-generated text. *arXiv preprint arXiv:2010.03070*. <https://doi.org/10.48550/arXiv.2010.03070>
- [10] Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis: Journal of English Language, Literature, and Teaching*, 5(1), 65-78. <http://dx.doi.org/10.31002/metathesis.v5i1.3519>
- [11] Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*, 6(1), 75. <https://doi.org/10.1038/s41746-023-00819-6>
- [12] GPTZero. (2023). [Web-based software]. <https://gptzero.me/>
- [13] Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., ... & Wu, Y. (2023). How close is chatgpt to human experts? comparison corpus, evaluation, and detection. *arXiv preprint arXiv:2301.07597*. <https://doi.org/10.48550/arXiv.2301.07597>
- [14] Ippolito, D., Duckworth, D., Callison-Burch, C., & Eck, D. (2019). Automatic detection of generated text is easiest when humans are fooled. *arXiv preprint arXiv:1911.00650*. <http://dx.doi.org/10.18653/v1/2020.acl-main.164>
- [15] Khalil, M., & Er, E. (2023). Will ChatGPT get you caught? Rethinking of plagiarism detection. *arXiv preprint arXiv:2302.04335*. <https://doi.org/10.48550/arXiv.2302.04335>
- [16] Krishna, K., Song, Y., Karpinska, M., Wieting, J., & Iyyer, M. (2023). Paraphrasing evades detectors of ai-generated text, but retrieval is an effective defense. *arXiv preprint arXiv:2303.13408*. <https://doi.org/10.48550/arXiv.2303.13408>
- [17] Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *arXiv preprint arXiv:2304.02819*. <https://doi.org/10.48550/arXiv.2304.02819>
- [18] Pegoraro, A., Kumari, K., Fereidooni, H., & Sadeghi, A. R. (2023). To ChatGPT, or not to ChatGPT: That is the question!. *arXiv preprint arXiv:2304.01487*. <https://doi.org/10.48550/arXiv.2304.01487>
- [19] Perkins, M., Roe, J., Postma, D., McGaughan, J., & Hickerson, D. (2023). Game of Tones: Faculty detection of GPT-4 generated content in university assessments. *arXiv preprint arXiv:2305.18081*. <https://doi.org/10.48550/arXiv.2305.18081>

- [20] Sapling. (2023). AI Detector. [Web-based software]. <https://sapling.ai/ai-content-detector/>
- [21] Turing, A. M. (2012). Computing machinery and intelligence (1950). *The Essential Turing: the Ideas That Gave Birth to the Computer Age*, 433-464.
- [22] Walters, W. H. (2023). The effectiveness of software designed to detect AI-generated writing: A comparison of 16 AI text detectors. *Open Information Science*, 7(1), 20220158. <https://doi.org/10.1515/opis-2022-0158>
- [23] Wang, J., Liu, S., Xie, X., & Li, Y. (2023). Evaluating AIGC Detectors on Code Content. *arXiv preprint arXiv:2304.05193*. <https://doi.org/10.48550/arXiv.2304.05193>
- [24] Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., ... & Waddington, L. (2023). Testing of detection tools for AI-generated text. *arXiv preprint arXiv:2306.15666*. <https://doi.org/10.48550/arXiv.2306.15666>
- [25] Writer. (2023). AI Content Detector. [Web-based software]. <https://writer.com/ai-content-detector/>
- [26] Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122-1136. <https://doi.org/10.1109/JAS.2023.123618>
- [27] Yan, D., Fauss, M., Hao, J., & Cui, W. (2023). Detection of AI-generated essays in writing assessment. *Psychological Testing and Assessment Modeling*, 65(2), 125-144.
- [28] Zemčík, M. T. (2019). A brief history of chatbots. *DEStech Transactions on Computer Science and Engineering*, 10. <http://dx.doi.org/10.12783/dtscse/aicacae2019/31439>
- [29] ZeroGPT. (2023). AI Text Detector. [Web-based software]. <https://www.zerogpt.com/>

## Appendix A

### Student Interview on the Use of AI Software for School Assignments

1. Have you ever used AI software for writing before? (If so please explain)
2. What type of AI software did you use for this project? (ex: ChatGPT)
3. On which assignments did you use AI software?
  - Population problems
  - Money issues
  - Genetically modified foods
  - Changes to the Japanese Constitution
4. How did you use the software? Please explain your methods.
5. What did you think of the experience of writing with the AI tool?
6. How did you feel about using AI-generated writing for your in-class discussions?

7. What is your overall feeling about students using AI software for schoolwork?

8. Do you have any more thoughts to add that did not come up from the above questions?

## Appendix B

**Do you think that there is a population problem? If Yes – How can we fix the problem? If No - Explain why you think that there is not a problem.**

Yes, I think that there is a population problem. I believe that declining birth rates, including in Japan and other major advanced countries, are a significant problem. To address the issue of declining birth rates, several methods can be considered:

1. Improve working conditions and promote work-life balance: Promoting workstyle reforms, providing flexible working conditions, and enhancing parental leave systems can make it easier for individuals to balance work and family life. Creating an environment where people feel secure to have children is essential.

2. Enhance economic support measures: Expanding economic support measures such as child allowances and childcare leave benefits can alleviate the financial burden of raising children. Policies to reduce the costs of education and healthcare are also worth considering.

3. Develop childcare environments: It is crucial to establish a comprehensive infrastructure to support child-rearing, including improving childcare facilities, enhancing the quality of education, and building community networks for childcare support. This ensures that a nurturing environment is available for raising children, fostering a sense of security for the future.

4. Flexible immigration policies: In countries experiencing declining birth rates, actively accepting foreign workers and immigrants can help mitigate labor shortages and population decline. Flexibility in immigration policies and achieving social harmony in accepting immigrants are important considerations.

## Appendix C

**How would revising the Japanese Constitution change our future?**

If the Japanese Constitution were to be revised, I believe that Japan would be under the Liberal Democratic Party of Japan. It is because we have experienced the failure of Opposition Coalition Party in

2009, and the LDP submitted Japan Constitution Amendment Draft in April 2012 (Miyamoto, Ueno and others, 2022). Article 9 would be a particularly contentious topic of discussion about revising Japanese Constitution. If it were to be changed, it would resolve the debate on the unconstitutionality of the Self-Defense Forces. However, it may not be good for relationship with countries around Japan. According to the LDP (2023), the principles of pacifism would remain unchanged, but other countries may regard as a willingness for war. It is important to maintain friendly relations with other nations, because Japan does not have a high self-sufficiency rate for food and enough resources. For this reason, I propose explicitly stating that the constitutional revision is for the purpose of self-defense and continuing peacekeeping activities.

Additionally, the Article 24 is one of the factors preventing the recognition of same-sex marriage in Japan. Therefore, revising Japanese Constitution will lead to the recognition of same-sex marriage.

Source :

Kenichi Miyamoto, Shigeki Uno, Noriho Urabe, Syogo Noda, Toru Morotomi, Yoichi Torihata, Hiroyuki Mori, Keiichi Yamazaki. [Politics and Economy for high school students -a newly revised edition-].2022.

Liberal Democratic Party of Japan (Taro Aso, Sinzo Abe, Yasuo Hukuda, Yoshiro Mori, others.) [Japan Constitution Amendment Draft]. 2012.

[https://storage.jimin.jp/pdf/news/policy/130250\\_1.pdf](https://storage.jimin.jp/pdf/news/policy/130250_1.pdf)

The official website of Liberal Democratic Party of Japan. [The four things we want to change]. Last time I read this website is 2023/06/27.

<https://www.jimin.jp/kenpou/proposal/>