A Quantitative Evaluation of Turnitin from an L2 Science and Engineering Perspective

Kathryn Oghigian (oghigian@gmail.com)
Waseda University, Japan
Michael Rayner (mikeyrayner@gmail.com)
Waseda University, Japan
Kiyomi Chujo (chuujou.kiyomi@nihon-u.ac.jp)
Nihon University, Japan

Abstract
The purpose of this paper is to investigate the functionality and accuracy of Turnitin results as applied to 68 science and engineering research papers, and the potential use of the software in a second language context. Results showed Turnitin found "similar matching" in 99% of papers; however, an analysis eliminating false positives and categorizing actual plagiarism events as *outright*, *paraphrase* and *patchwork plagiarism*, or *stealing an apt term* showed only 29% featured plagiarized material, and in most cases, evidence suggested no intent to deceive. Findings indicate that Turnitin can be useful, particularly as a pedagogical rather than policing tool, but "similarity" percentages can be misleading and careful user evaluation of the entire paper shown with flagged highlighting is necessary in order to fairly assess student intent.

Keywords: plagiarism detection, originality, second language learning, Turnitin

Introduction

From a review of the huge number of studies in the literature, it is immediately apparent that plagiarism and the use of plagiarism detection software are challenging issues. A great deal of research has been done on plagiarism (for an excellent overview, see Flowerdew & Li, 2007:162); however, it is also clear that there is no agreed upon definition or nomenclature, particularly when applied to the reproduction of text, as opposed to ideas (see Pecorari, 2001). Suggested terms include textual plagiarism (Flowerdew & Li, 2007), textual borrowing (Shi, 2006), textual and prototypical plagiarism (Pecorari, 2001), and plagiphrasing (Whitaker, 1993). Other researchers have suggested that plagiarism is a Western standard that dates back to the appearance of the printing press in the Middle Ages (Flowerdew & Li, 2007) and is not applicable in other cultures, particularly those in Asia and the Middle East (Click, 2012; Sowden, 2005). The use of plagiarism software for student writing is also an issue. Turnitin (http://turnitin.com/), arguably the most commonly used plagiarism detection software, has been banned from several universities, most notably Yale, Harvard, Princeton (*The Daily Princetonian*, 2006; Hanrahan, 2008; Bretag & Mahmud, 2009), and others, and its

mandatory use has been legally challenged in other academic institutions (*CBC*, 2004; *The McGill Daily*, 2005).

Second Language Learning and Plagiarism

In a second language (L2) context, there seem to be several important issues concerning plagiarism. These include whether students fully grasp the concepts of similarity (i.e., how strings of words can be matched as "similar" to words in a database by software), originality and plagiarism; what role cultural expectations may play; and if the practice of plagiarism may simply be a developmental issue, which disappears as students increase their productive vocabulary and learn how to take better notes, summarize, paraphrase and quote sources. In their detailed qualitative examination of Turnitin through student questionnaire responses on its use, Bensal, Miraflores and Tan (2013) explored the first of these issues and concluded that the students in their study tended to miss the point of originality; in other words, that they focused their efforts on "rephrasing to evade high percentages of Similarity Index ratings more than trying to explain the real concepts in the paper guided by a real sense of academic integrity" (p. 17). Some international students may not intend to deceive, but rather engage in various forms of plagiarism because they believe memorization and recitation to be acceptable and valuable, or they may have difficulty understanding that someone can own an idea (Sutherland-Smith, 2005; Trudeau & Sevier, undated). Shi (2006) investigated plagiarism ("textual appropriation") and found that students did not understand what needed citing and what did not; and nonwestern students viewed the idea of plagiarism as "foreign and unacceptable" (p. 264). Jones and Freeman (2003) describe a "benign learning process" in which students learn to copy format, segments, and phrases, resulting in the perception of copying as a valid form of learning. L2 students are often given "sentence templates," or taught to search out clusters in corpora to incorporate into writing (Hunston, 2002), and the introduction of corpus pedagogy to find and use common expressions or formulaic writing may contribute to learner uncertainty, i.e., confusion between what is common usage and allowable, and what is unique, and not. Flowerdew and Li (2007) summarize the use of plagiarism as a student survival strategy based on the belief that some copying is acceptable to combat task overload and pressure to pass assessments, and to compensate for a lack of confidence in using the target language (p. 169).

These studies offer a glimpse into why students might plagiarize; however, there are limited studies on evaluating the functionality and accuracy of Turnitin in an L2 context. Stapleton (2012) evaluated whether or not Turnitin was a deterrent for plagiarizing in a study on writing produced by L1 graduate students. He found that students who were aware their work would be checked had lower percentages, but also noted that the software was not necessarily accurate and should be used with caution. Walker (2010) used Turnitin to assess the frequency, nature and extent of plagiarism in university business student writing, but did not assess its accuracy or provide suggestions for more effective use.

Purpose of the Study

Turnitin was recently adopted by a Japanese university and made available to English academic writing teachers (two of the authors). The purpose of this paper is to investigate how the software could be best employed in an L2 science and engineering context by evaluating the functionality and accuracy of the program; and suggesting modifications that could make it a more effective tool for L2 students and teachers.

In this study, two important assumptions are made. First, it is recognized that Turnitin was not created as a teaching tool for L2 writing students. This investigation is not a criticism of the software but rather is an excursion into how it functions in an L2 environment and how it might be adapted to be more effective. Secondly, Turnitin was originally created and marketed as a plagiarism detection software (Barrie & Presti, 2000), and in the perceptions of many users remains so (Shi, 2006; Stapleton, 2012; Vie, undated). Turnitin provides an initial percentage of "similar" or "matching" text and flags questionable ("similar") passages; this study is an evaluation of this percentage and specific flagged strings of words in order to more clearly understand whether or these results can, in fact, be viewed as plagiarism.

Method

Operational Definitions

Turnitin is only able to identify language in the form of sequential words or strings that match against its database through its use of algorithms, however plagiarism in the sense of the use of another's *ideas* without citation, is not measured or evaluated. Thus an evaluation of the software is only possible in terms of how effectively it identifies to what extent students use identical phrasing to source texts or other students' work.

Participants

Research papers written by 68 third and fourth year undergraduate students as the final task of an academic writing course were screened for plagiarism using Turnitin. The students belonged to various departments in the faculty of science and engineering. The students had a wide range of English language abilities based on teacher observation, and self-reported Test of English for International Communication (TOEIC) scores ranging from 390 to 950. They were enrolled in a weekly 90 minute, one semester (14 week) elective writing course with a goal of producing a 2,000-word research paper on a topic of their choice in their technical fields, formatted in IEEE style (http://www.ieee.org/documents/stylemanual.pdf). Students were given instruction in research writing, including how to cite and reference sources and how to take notes in their own words and quote, paraphrase and summarize. University policy governing plagiarism was provided both in the L1 (Japanese) and L2 (English), and students were repeatedly reminded "not to copy" because "plagiarized papers will fail." Participants signed a consent form to allow their work to be used for research purposes.

Procedure

The university provides Turnitin for teachers' use, and the interface is found on the university network. To access Turnitin, a teacher creates a folder into which each student submits a paper. A settings option for Turnitin is available to the teacher so that all papers will be automatically checked when uploaded by the students. A percentage denoting the "Plagiarism Detection Results" and a coloured box representing the scale of the problem appear next to each student's name once the paper completes its run through the program. In Figure 1, the first result shows 28% with a yellow code; subsequent results are shown with various percentages and green codes. The figure is truncated to conceal identifying information about the students.

Figure 1. A partial class list with colour-coded "Plagiarism Detection Results".

Functionality

For the first analysis, the original colour-coded "plagiarism detection" percentage shown in the folder class list was noted for each student paper. Turnitin filters allow users to identify and exclude quotes and bibliography from the "Plagiarism Detection Results" percentage, and these filters were then activated so that quotes and bibliography [hereafter referred to as references] would be excluded. The new percentage was noted. (This had to be done manually for each paper because of this particular system set-up.)

Accuracy

By clicking on the underlined percentage next to "plagiarism detection results" in the class list, a document viewer is loaded. This shows the entire paper along with a different

set of numbered, colour-coded highlighting of "similar" text corresponding to sources in a list on the right having the same text. (An example is shown in Figure 2.) In this document view of Turnitin, each instance of highlighted phrasing was captured using screenshots and was labelled with a student number S1 through S68. Multiple screenshots for one paper were given subset labels of a, b, and c, so for example, four screenshots for Student 1 would appear as 1a, 1b, 1c and 1d. This allowed co-researchers to analyze the data while protecting student anonymity.

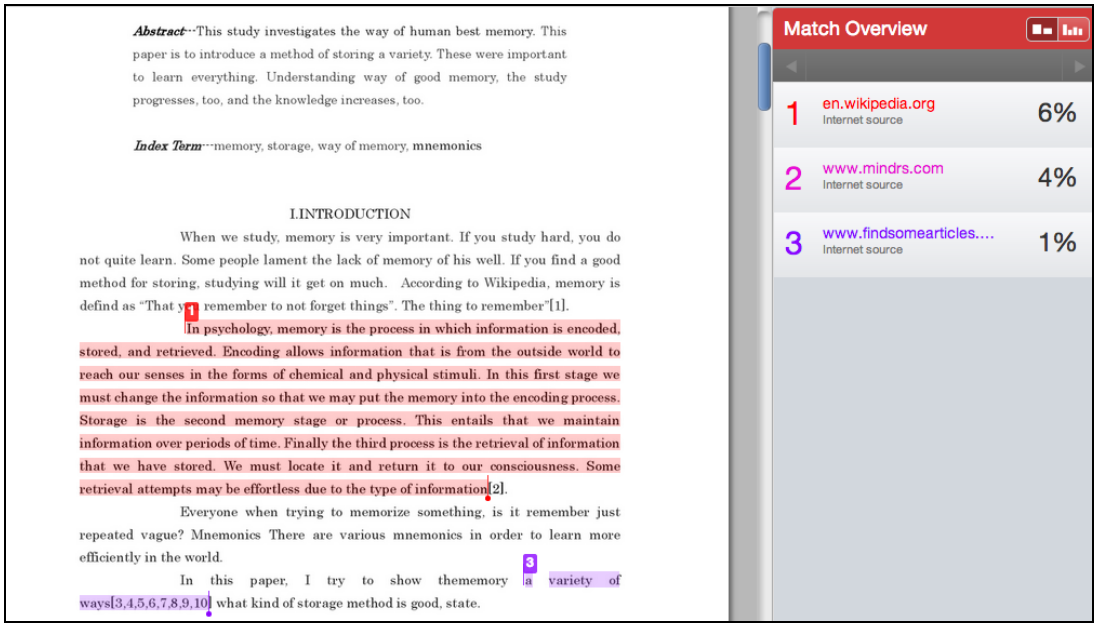


Figure 2. The document view of a paper. The red highlighting with the number 1 on the left is matched as exact text to the first source on the right (Wikipedia). Different colours represent different sources, but additional text throughout the paper can be linked back to the same number.

Each paper was then carefully examined for false positives. These included (a) page numbers, (b) affiliations, (c) references or portions of references and/or (d) quotes that were flagged even though the filter was applied, (e) mathematical formulae, (f) the cited reproduction of tables or figures, (g) the cited captions of tables or figures, (h) phrases specifically taught to students, and (i) common expressions.

Application to L2 pedagogy

Finally, the detailed document view of the program was used to classify the remaining instances of highlighted phrasing based on four levels and definitions of plagiarism from Hamp-Lyons and Courter (1984:161-166): *outright copying* (copying whole sentences and paragraphs); *paraphrase plagiarism* (replacing a few words and phrases but copying the general sentence structure and grammar); *patchwork plagiarism* (taking various phrases from one or multiple sources and putting them together in a different way); and *stealing an apt term* (copying a pithy or memorable short phrase).

Results and Discussion

Functionality

The Turnitin colour-coding system divides the results into five colour-coded categories, based on the percentage of text having "similarity" to text in the software database. Blue indicates no matching text, green is for a similarity percentage of 24% or less, yellow is for 25% to 49%, orange is for 50% to 74%, and red is for 75% to 100% matching text (see Turnitin, Instructor User Manual, Ch. 2, p. 49). The original colour-coded percentage is the first screen viewed by a Turnitin user (in Figure 1), and the number of papers falling into each category are shown in the first column of each pair of columns in Figure 3 (N=68). As noted earlier, the Turnitin program on this university system requires a teacher to manually exclude quotes and references for each paper. The number of papers in each category after this exclusion are shown in the second column of each pair of columns in this figure.

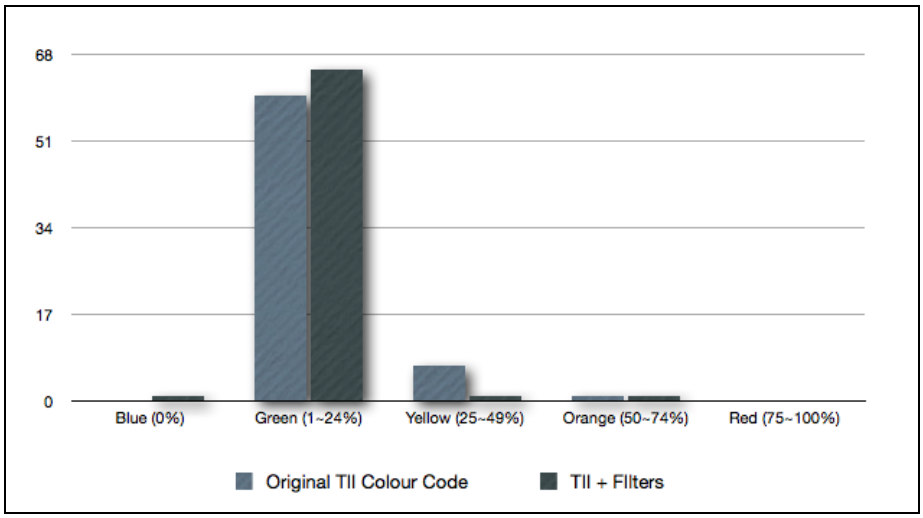


Figure 3. Number of research papers categorized into Turnitin colour codes.

One striking result was that, according to the Turnitin analysis, none of the 68 students were able to write a completely original paper, as shown by no results in the first blue column on the left. There were 60 papers initially coded as green (one "similar" or possibly plagiarized word to 24% of the text), seven papers coded as yellow (25% to 49% possibly plagiarized text), and one paper as orange (50% to 74%). However, after the reference and quotes filters had been applied, the results showed one blue (original) paper, 65 green, one yellow, one orange and no red.

Also notable is that a student's paper could match up to 24% of its content and still receive a green code. This is important because an instructor, particularly a busy instructor with a great many papers to assess, may simply look at the green code on a class list (such as the one in Figure 1) and dismiss these as not requiring further attention. Once references and quotes are filtered from the results, the Turnitin percentage may also be low enough that a teacher might dismiss it when in fact there are problems that need to be addressed. In the paper shown in Figure 2, the Turnitin result was originally given as

16% (green), and post-filters as 11% (green). A closer examination of this "acceptable level" showed that the student plagiarized the entire second paragraph of the introduction (shown in Figure 2) and the majority of another paragraph (not shown). This is easily seen in the full document view by the uninterrupted string of highlighted words, and was *outright plagiarism*, although a citation was provided; the document view function is clearly helpful to the teacher and would be helpful to students who had access. With regard to colour code function, the green code may serve as an erroneous visual cue (green means "go") that plagiarism is either not present or exists in low (and therefore "acceptable") levels, thus this software program used for policing is clearly not always effective. Furthermore, the percentage does not indicate the number of words in a copied (or "similar") string, so unless the instructor looks at the entire paper in the document view, portions of *outright* copying would not be identified. Therefore, the data here indicate that it is important not to rely only on initial percentages, but to use the document viewer.

Accuracy

While doing this analysis, it was immediately clear that the software produced a number of false positives. These were identified manually and papers including false positives were counted. The number of papers and categories of these false positives are shown in Figure 4. The false positives appeared to be a consequence of two factors concerned with the functions of the software: the lack of reliability of the filters which are provided, and the absence of other desirable filters from the software design.

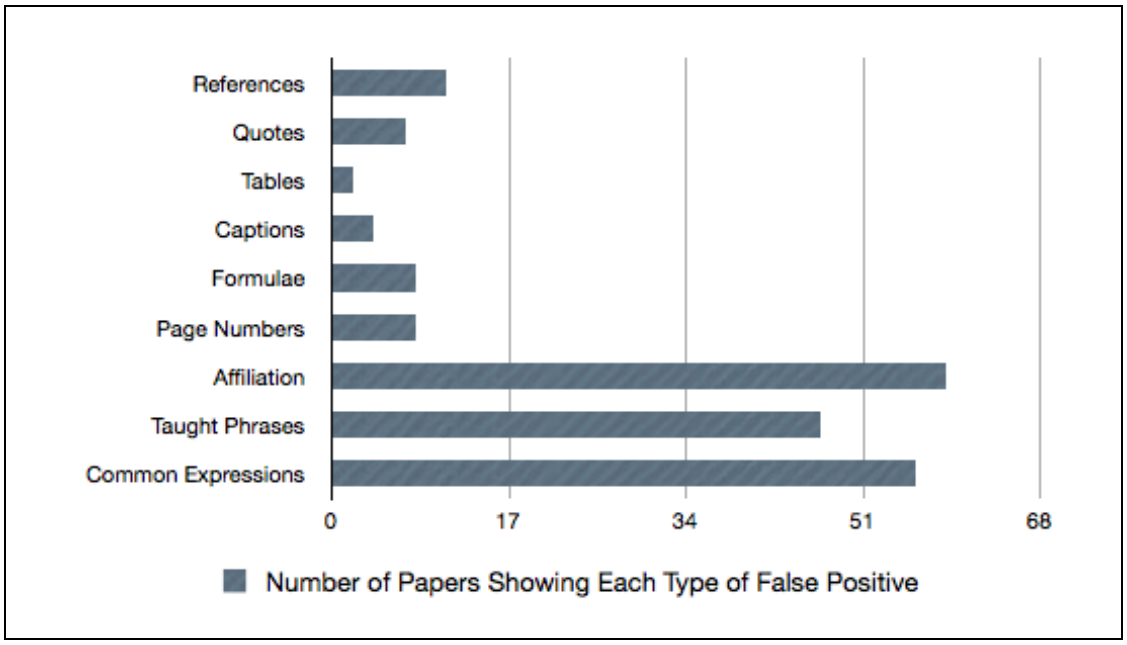


Figure 4. Categories of false positives shown in the number of papers (N=68).

Unreliable filters

Flagged references occurred in 18 (16%) of the post-filtered papers in spite of activating the filter. An example is shown in Appendix Figure A1. Quotes were similarly flagged in seven papers (10%). In two papers, the flagged quotes were long, i.e., 40 words or more, and were indented (similar to APA style) but without quotation marks. In one case, two quotes were shown in the same screenshot, but Turnitin nevertheless flagged one even with this "no quote" filter in place. Both appear to be identical in the use of quotation marks, spacing, and the incorrect placement of the citation inside the ending quotation mark. (See Appendix, Figure A2.) It is clear, therefore, that the filters can be unreliable.

Missing filters

The software does not allow the user to filter out certain other common features of academic papers, shown in Figure 4, which produced a number of false positives. Turnitin flagged reproduced and appropriately cited tables in two papers, and seven captions in six papers. Students had been instructed to use "Figure/table reproduced from [citation]" and did so for all but one caption. In eight papers (12%), mathematical formulae were flagged as "matching". Mathematical strings comprising equations are typical in mathematics papers, and these strings were frequently flagged as "similar". (An example is provided in Appendix Figure A3). From a total of 68 papers, 56 used page numbers, and of these, page numbers in eight papers were flagged as "matching." (An example is provided in Appendix Figure A4). Additionally, Turnitin flagged the department, faculty and/or university in 59 papers (87%) (Appendix Figure A5). While this only showed up as 1% or 2% of the total, it nevertheless inflated the final "plagiarism" percentage, and perhaps more importantly, was a distraction to the instructor which required additional clicking (and therefore time) to eliminate.

Students in the course were taught various sentence templates, specific phrases and logical connecting or signaling devices for expressing the purpose of the paper, describing the research as secondary-sourced, presenting data, transitioning, hedging and making inferences. Turnitin flagged taught phrases in 47 papers (69%) (see Figure 4). Examples are shown in Figure 5.

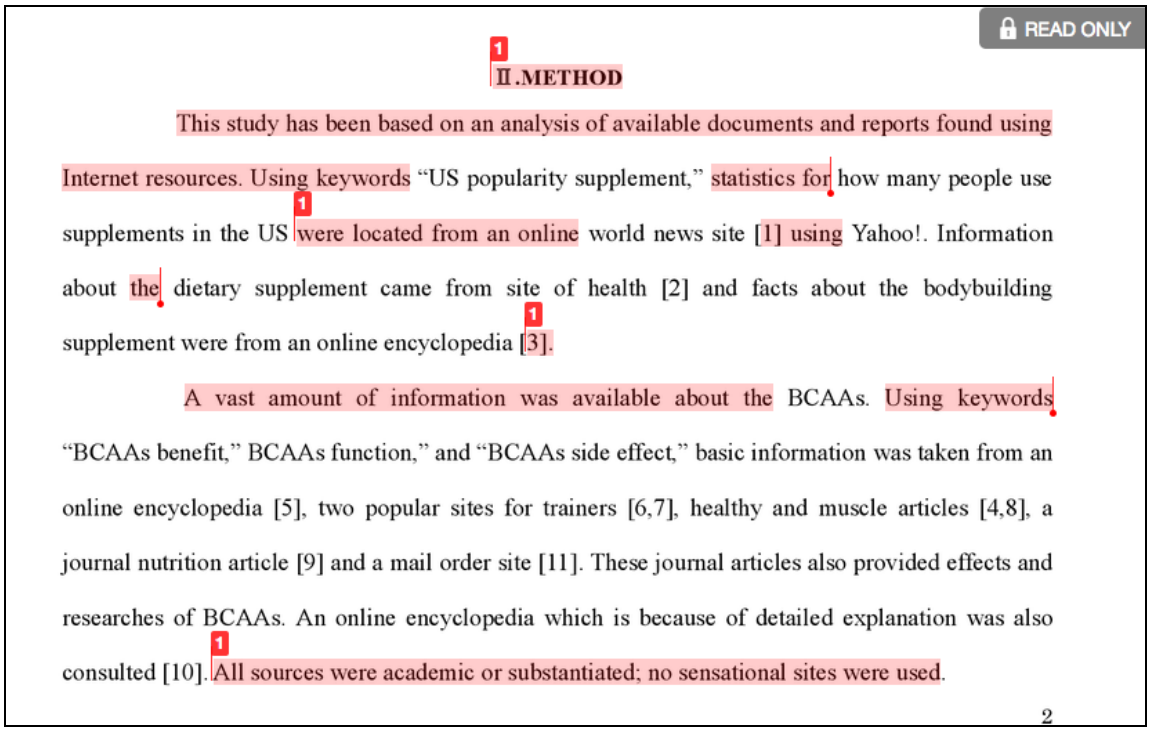


Figure 5. Sentence templates and fixed phrases flagged by Turnitin. In this case, the highlighting (shown as "1") for similarity was matched to phrases used in university papers from the same university by other students in the same course.

Common expressions, or prefabricated chunks of language, were the final major false positive identified. In more than 80% of the papers (N=56), Turnitin flagged fragments of text that were arguably benign combinations of words. Examples are shown in Table 1 and in Figure 6. In cases where there was uncertainty about the commonality of a phrase, a corpus search was done on the Corpus of Contemporary American English (COCA) (Davies 2008-) to determine if the expression was unique (zero or one occurrence) or not (multiple occurrences).

Table 1

Selected Examples of Common Expressions Flagged by Turnitin

	Common expressions
1	they are the largest animal in the world
2	evolutionary game theory…a…mathematical model…the…of…
3	data of…Ministry of Education, Culture, Sports, Science and Technology in
4	Waseda.jp Abstract This study investigates the causes of

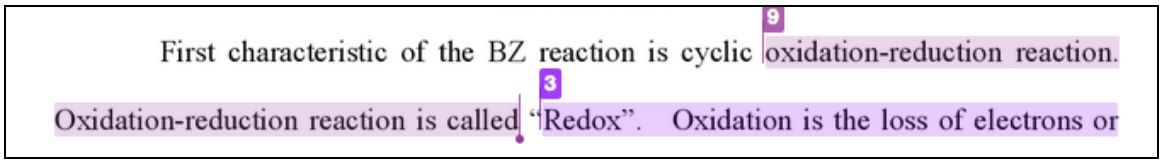


Figure 6. Common fragments that were flagged by Turnitin.

One might argue that it is possible to select the "exclude small matches" feature of Turnitin; this instructs the software to only to identify only long strings of text. However, use of this feature with the papers used in the present study would have eliminated shorter plagiarized segments such as the *broken safety culture at NASA* and the other examples shown in Table 2.

Table 2
Examples of Short Plagiarized Word Strings Flagged by Turnitin

1	broken safety culture at NASA
2	have operated in a challenging and [furious] environment
3	The green bars show uncertain estimates.
4	Challenger broke apart 73 seconds into its flight

To determine the accuracy of the software, after the false positives had been identified and excluded, a new percentage was calculated by dividing the number of actual plagiarized words in each student paper by the total number of words and multiplying by 100. Figure 7 shows a comparison of the initial (filtered) Turnitin colour-coded results with the results after the false positives were excluded. The data in the figure shows that although Turnitin had originally determined that only one student was able to write a paper with no plagiarism issues, there were 46 of these post instructor analysis. Turnitin rated 65 papers green; the instructor analysis yielded 18. Both Turnitin and the instructor identified one paper in each of the yellow and orange rating categories, and none in red. These results suggest that the software is inaccurate because of unreliable filter function (for quotes and references) and missing filters (for tables, captions, mathematical formulae, page numbers, affiliations, taught phrases and common expressions).

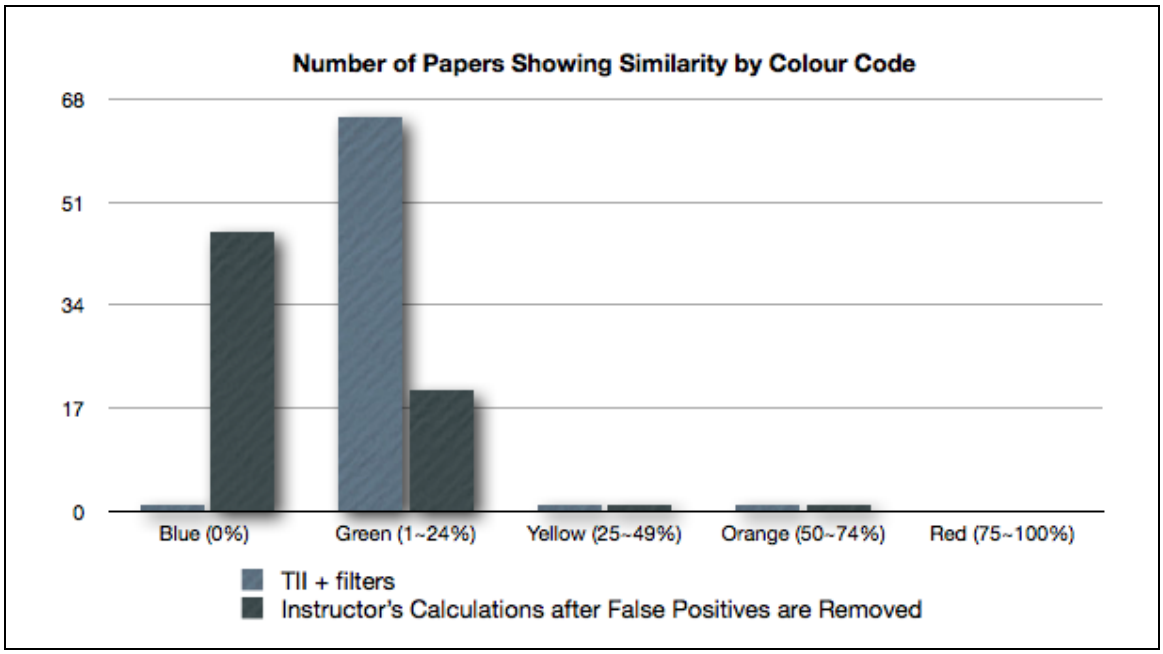


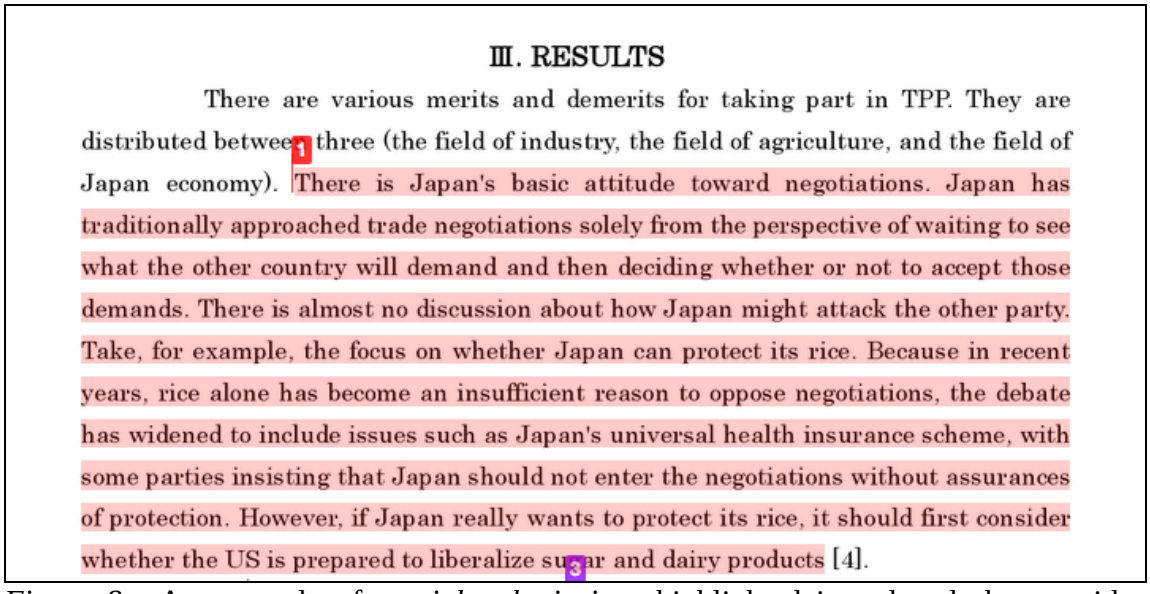
Figure 7. A comparison of Turnitin and instructor's results.

Definitions

Definitions were not initially counted as a false positive, but in reviewing papers, it was clear that, particularly in explaining technical science and engineering concepts, the presentation of definitions was an issue. In this course, students were taught to present definitions by providing a source, for example: "According to <source>, a <term> is defined as "…<definition>" [citation]. Although many students followed this pattern, other definitions, which can reasonably considered to be general and common, appeared in eleven student papers and were flagged by Turnitin but allowed by the instructor. An example of a general or common definition is shown in the Appendix (Figure A6).

Plagiarism

After eliminating false positives, actual plagiarism was found in 20 (29%) student papers (18 green, one yellow, and one orange), in contrast to the initial 68 (100%) of possible plagiarism indicated by Turnitin. Each Turnitin flagged string was next categorized by the authors as *outright copying* (shortened to *outright*), *paraphrase plagiarism* (*para*), *patchwork plagiarism* (*patch*), or *stealing an apt term (apt)*. Examples of each type are shown in Figures 8 through 11, respectively. Plagiarism was viewed as *outright* if a few words were added before or after the main string of text, such as "Therefore, <plagiarized string>" or a misspelling or minor word change interrupted the plagiarized string. Although the Turnitin colour code can be misleading, and the results can be inaccurate, the coloured highlighting of possibly plagiarized material is helpful to teachers when used in document view. Teachers are able to examine these strings to determine if there was an attempt to paraphrase by the student, and/or if a citation was given. This information is useful in directing students' attention to either clearing up a misunderstanding, or in identifying the need for more quoting, paraphrasing and summarizing practice.



Ⅲ. RESULTS

There are various merits and demerits for taking part in TPP. They are distributed between three (the field of industry, the field of agriculture, and the field of Japan economy). There is Japan's basic attitude toward negotiations. Japan has traditionally approached trade negotiations solely from the perspective of waiting to see what the other country will demand and then deciding whether or not to accept those demands. There is almost no discussion about how Japan might attack the other party. Take, for example, the focus on whether Japan can protect its rice. Because in recent years, rice alone has become an insufficient reason to oppose negotiations, the debate has widened to include issues such as Japan's universal health insurance scheme, with some parties insisting that Japan should not enter the negotiations without assurances of protection. However, if Japan really wants to protect its rice, it should first consider whether the US is prepared to liberalize sugar and dairy products [4].

Figure 8. An example of *outright plagiarism*, highlighted in red and shown with a citation.

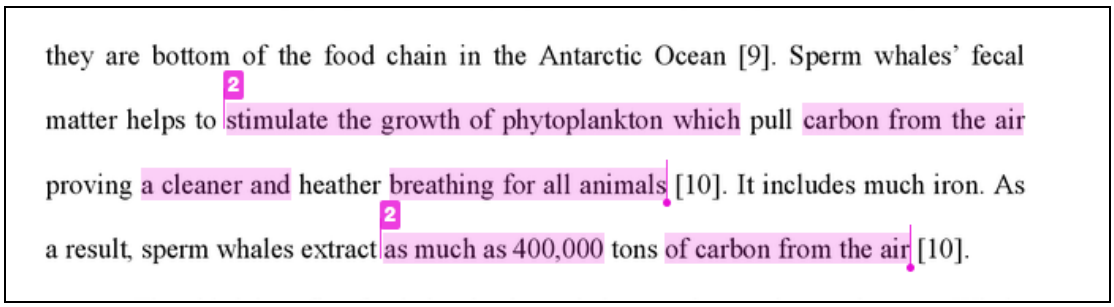

they are bottom of the food chain in the Antarctic Ocean [9]. Sperm whales' fecal matter helps to stimulate the growth of phytoplankton which pull carbon from the air proving a cleaner and heather breathing for all animals [10]. It includes much iron. As a result, sperm whales extract as much as 400,000 tons of carbon from the air [10].

Figure 9. An example of *paraphrase plagiarism*, with all phrases from the same source; citations are given.

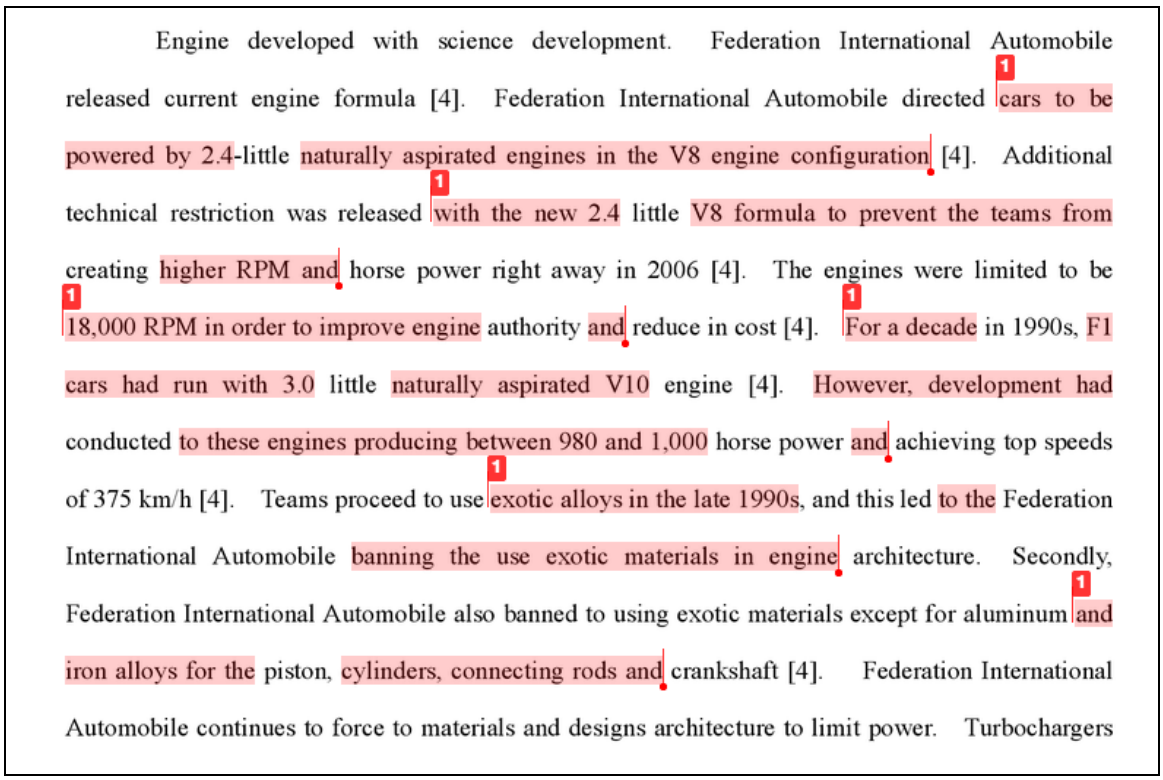

Engine developed with science development. Federation International Automobile released current engine formula [4]. Federation International Automobile directed cars to be powered by 2.4-little naturally aspirated engines in the V8 engine configuration [4]. Additional technical restriction was released with the new 2.4 little V8 formula to prevent the teams from creating higher RPM and horse power right away in 2006 [4]. The engines were limited to be 18,000 RPM in order to improve engine authority and reduce in cost [4]. For a decade in 1990s, F1 cars had run with 3.0 little naturally aspirated V10 engine [4]. However, development had conducted to these engines producing between 980 and 1,000 horse power and achieving top speeds of 375 km/h [4]. Teams proceed to use exotic alloys in the late 1990s, and this led to the Federation International Automobile banning the use exotic materials in engine architecture. Secondly, Federation International Automobile also banned to using exotic materials except for aluminum and iron alloys for the piston, cylinders, connecting rods and crankshaft [4]. Federation International Automobile continues to force to materials and designs architecture to limit power. Turbochargers

Figure 10. An example of *patchwork plagiarism*, all originating from the same source "1" which was Wikipedia; citations are given.

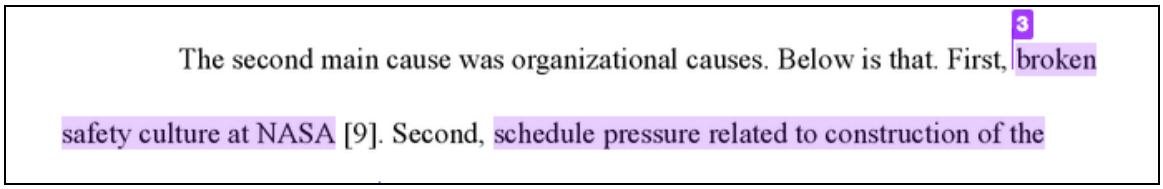

The second main cause was organizational causes. Below is that. First, broken safety culture at NASA [9]. Second, schedule pressure related to construction of the

Figure 11. An example of *stealing an apt term* ("broken safety culture at NASA").

As shown in Figure 12, of 20 papers, eight papers contained all four types of plagiarism; four contained only *outright*; and eight contained *para*, *patch* and *apt*. This suggests that of the 68 students, only 20 had difficulty with plagiarism--- four who used *outright* and 16 who attempted to incorporate phrases into their own writing. It should be noted that in

almost all cases, students provided a citation, almost always after each sentence even if all of the sentences in one paragraph repeated the same citation. *Para*, *patch* and *apt* were combined and contrasted with *outright* because the former show an attempt to paraphrase, whereas *outright* does not.

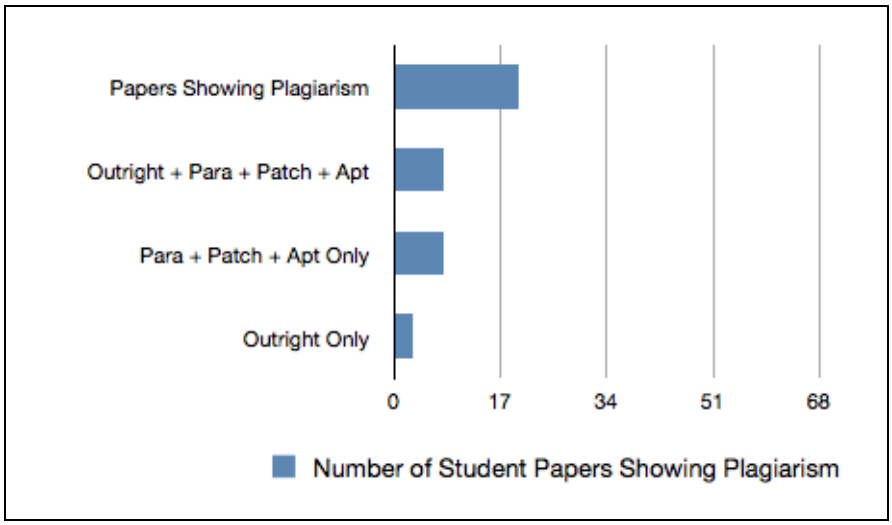


Figure 12. Number of student papers containing types of plagiarism.

Once the types of plagiarism were determined for these 20 papers, the amount of plagiarized content was then examined. Nearly all show low percentages of plagiarized material (6% or under), including two of those which contain only *outright plagiarism*. This suggests that the majority of these students attempted to paraphrase and the problematic content was minimal. Two papers with higher percentages contained all four types of plagiarism, one at 14%, and the other at 52% (rated orange by Turnitin); the use of citations and *para* and *patch* also suggest an attempt to paraphrase. The remaining two papers had 20% and 35% *outright plagiarism* content. Turnitin rated these as green and yellow respectively.

These results suggest that students understood that plagiarism was not acceptable and that citations and references were required for all facts or ideas originating from another source; however, it also seems clear that despite practice in note taking, summarizing, paraphrasing and quoting, many students struggled with these skills. For example, in Figure A3, the student attempts to paraphrase a mathematical proof and uses quotation marks and a citation, suggesting that although he is not clear on which strategy to use, he uses them all in order to avoid an accusation of plagiarism. The students who had copied full sentences and whole paragraphs were lower level proficiency students, and this also suggests that the use of copying is a developmental issue.

Conclusion

Turnitin was created as a tool for L1 writing; however, because Turnitin is one of the few plagiarism detection programs available, it has also been adopted by L2 writing teachers.

In this analysis, it was found that it can be very useful when used in document view, but initial colour-coded percentages can be misleading because of inaccuracies.

Turnitin might be much more effective if used as a pedagogical rather than a policing tool. When used in the document view, a student or teacher can quickly and easily see strings of similar or plagiarized words to inform teaching and learning. Although Turnitin does not definitively identify something as plagiarism or distinguish between types of plagiarism, features such as the density of flagged text and the number of sources from which paragraphs, sentences or phrases are "borrowed" can be determined relatively easily due to the manner in which text is highlighted and annotated in the results window of the program. The information gleaned from this analysis could be used to address the cause of the plagiarism, e.g., not understanding the concept of originality, a cultural expectation, confusion over what is common and unique, or the developmental level of the learner. For optimum learning opportunities, it would seem to be preferable for this analysis to be carried out by students individually before they submit papers; however, this would obviously require students to have access to the program.

Finally, in this age of social media, a person can coin an original or unique term or concept that can be picked up and repeated immediately and endlessly. Where is the line between original, similar and common? If we as educators cannot be clear, then how can students? Two of the largest categories of false positives were for specifically taught phrases and common expressions. Turnitin compares student writing to a corpus of "non-allowed" language. It would be useful in an L2 context if software could also compare student writing to an "allowable" corpus, so that common phrases would not be flagged. In the meantime, students can be taught to check phrases in a corpus such as COCA. Clearly, Turnitin and similar programs can be useful but require a careful review of all results.

References

Barrie, J., and Presti, D. (2000). Digital plagiarism- The web giveth and the web shall taketh. *Journal of Medical Internet Research 2*(1):e6). doi: 10.2196/jmir.2.1.e6 Retrieved from http://www.jmir.org/2000/1/e6/

Bensal, E., Mirafloes, E., & Tan, N. (2013). Plagiarism: Shall we turn to Turnitin? *CALL-EJ, 14*(2), 2-22. Retrieved from http://callej.org/journal/14-2/Bensal_Miraflores_Tan_2013.pdf

Bretag, T. & Mahmud, S. (2009). Self-plagiarism or appropriate textual re-use? *Journal of Academic Ethics, 7*, 193-205. doi: 10.1007/s108055-009-9092-1

CBC (Canadian Broadcasting Corporation). (2004, January 16). McGill student wins fight over anti-cheating. Retrieved from http://web.archive.org/web/20070517033653/http://www.cbc.ca/news/story/2004/01/16/mcgill_turnitin030116.html

Click, A. (2012). Issues of plagiarism and academic honesty for second language students. *MELA Notes, 85*, 44-53. Retrieved from

http://www.jstor.org/discover/10.2307/23392491?uid=2129&uid=2&uid=70&uid=4&sid=21102844626417

Davies, M. (2008-). *The corpus of contemporary American English: 450 million words, 1990-present*. [Online corpus]. Retrieved from http://corpus.byu.edu/coca/.

Flowerdew, J. & Li, Y. (2007). Plagiarism and second language writing in an electronic age. *Annual Review of Applied Linguistics, 27,* 161-183. doi: 10.1017/S0267190508070086

Hamp-Lyons, L., & Courter, K. (1984*). Research matters*. Rowley Mass; Newbury House.

Hanrahan, M. (2008). Plagiarism, instruction and blogs. In T.S. Roberts (ed.): *Student Plagiarism in an Online World: Problems and Solutions* (pp. 183-193). NY: Information Science Reference.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Jones, A., & Freeman, T. (2003). Imitation, copying, and the use of models: Report writing in an introductory physics course. *IEEE Transactions of Professional Communication, 46*(3), 168–184.

Pecoari, D. (2001). Plagiarism and international students: How the English-speaking university responds. In D. Belcher & A. Hirvela (eds.): *Linking Literacies: Perspectives on L2 Reading-Writing Connections* (pp. 229-245). Ann Arbor, MI: The University of Michigan Press.

Shi, L. (2006). Cultural backgrounds and textual appropriation. *Language Awareness 15*(4), 264-282. doi: 10.2167/la406.0. Retrieved from http://www.tandfonline.com/doi/abs/10.2167/la406.0#.VaXKRMaqqko

Sowden, C. (2005). Plagiarism and the culture of multilingual students in higher education abroad. *English Language Teaching Journal, 59*(3), 226-233.

Stapleton, P. (2012). Gauging the effectiveness of anti-plagiarism software: An empirical study of second language graduate writers. *Journal of English for Academic Purposes 11*(2), 125-133. doi:10.1016/j.jeap.2011.10.003. Retrieved from http://www.sciencedirect.com/science/article/pii/S1475158511000749

Sutherland-Smith, W. (2005). Pandora's box: Academic perceptions of student plagiarism in writing. *Journal of English for Academic Purposes 4*, 83–95.

The Daily Princetonian. (2006, April 4). University opts not to 'Turnitin'. Retrieved from http://dailyprincetonian.com/news/2006/04/university-opts-not-to-turnitin/

The McGill Daily. (2005, December 2). Students: 2, Turnitin: 0. Senate committee rules in favour of student two years after she failed class for refusing to submit paper to Turnitin.com. Retrieved from http://web.archive.org/web/20070517093213/http://www.mcgilldaily.com/view.php?aid=4615

Trudeau, A. & Sevier, M. (undated). Paraphrasing for academic success. Powerpoint presentation, English Bridge Program, Simon Fraser University. Retrieved from www.sfu.ca/~msevier/paraphrasetrudeauseviernn.ppt

Turnitin Home Page. http://turnitin.com/

Turnitin Instructor User Manual. Retrieved from http://turnitin.com/en_us/training/instructor-training/instructor-user-manual

Vie, S. (undated). Turn it down, don't Turnitin: Resisting plagiarism detection services by talking about plagiarism rhetorically. Fort Lewis College website. Cited July 14, 2015. Retrieved from http://www2.bgsu.edu/departments/english/cconline/spring2013_special_issue/Vie/history.html

Walker, J. (2010). Measuring plagiarism: Researching what students do, not what they say they do. *Studies in Higher Education 35*(1), 41-59. doi: 10.1080/03075070902912994. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/03075070902912994#.VaXM5saqqko

Whitaker, E. (1993). A pedagogy to address plagiarism. *College Composition and Communication, 44*, 509–514.

Appendix 1. Additional Examples of False Positives Flagged by Turnitin

Figure A1. A screenshot of Turnitin showing false positives for references, even though this filter had been selected.

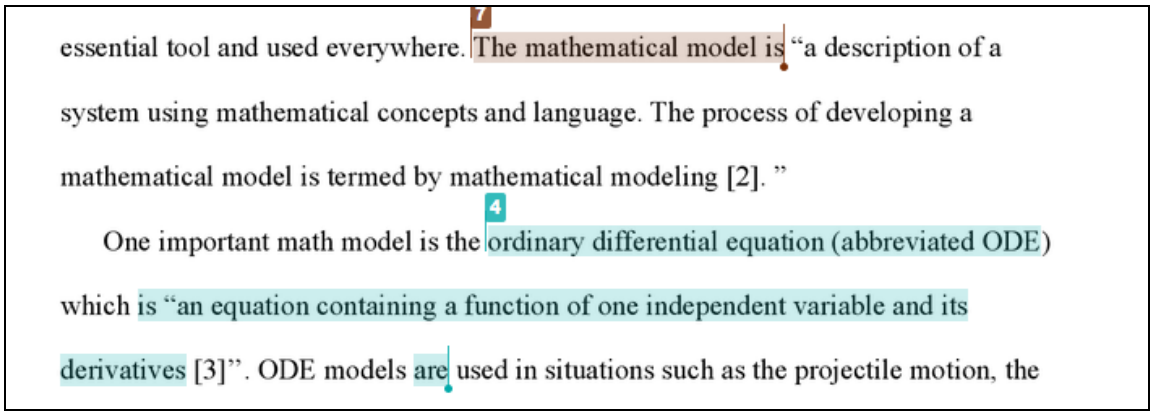


Figure A2. A screenshot of Turnitin showing one quote not flagged and a false positive for an identically-formatted second quote in the same paragraph, even though this filter had been selected.

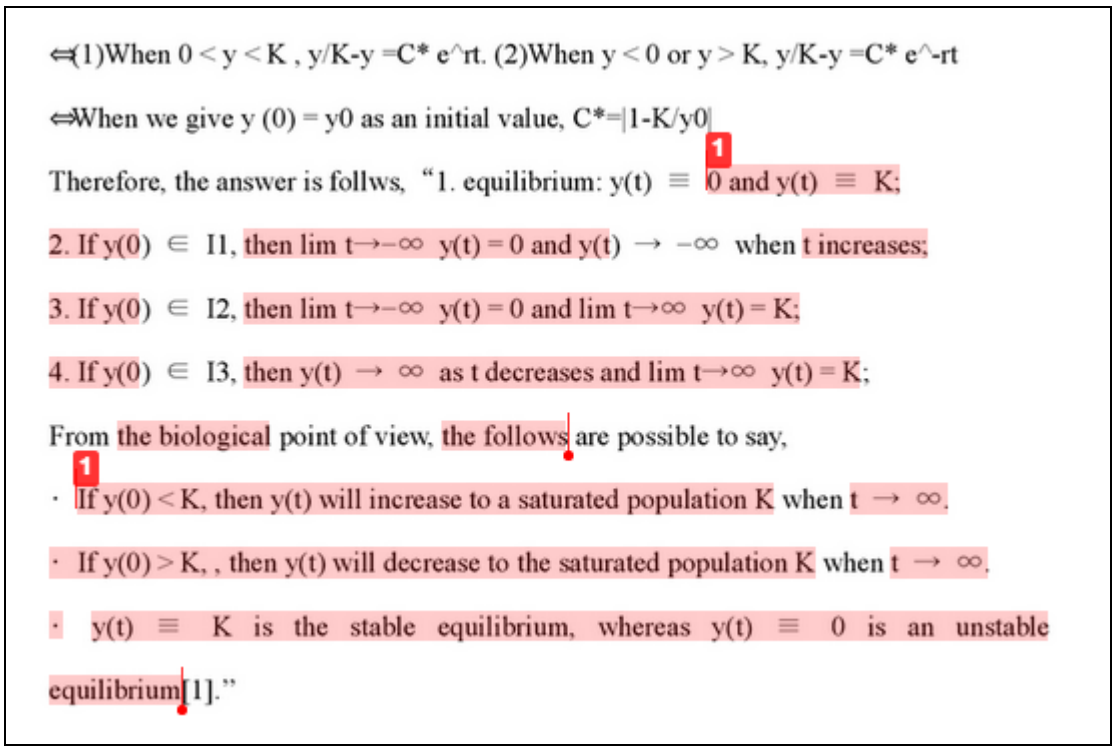


Figure A3. A screenshot of a flagged theorem. The student also provides a citation and quotation marks (albeit incorrectly formatted), suggesting no intent to plagiarize.

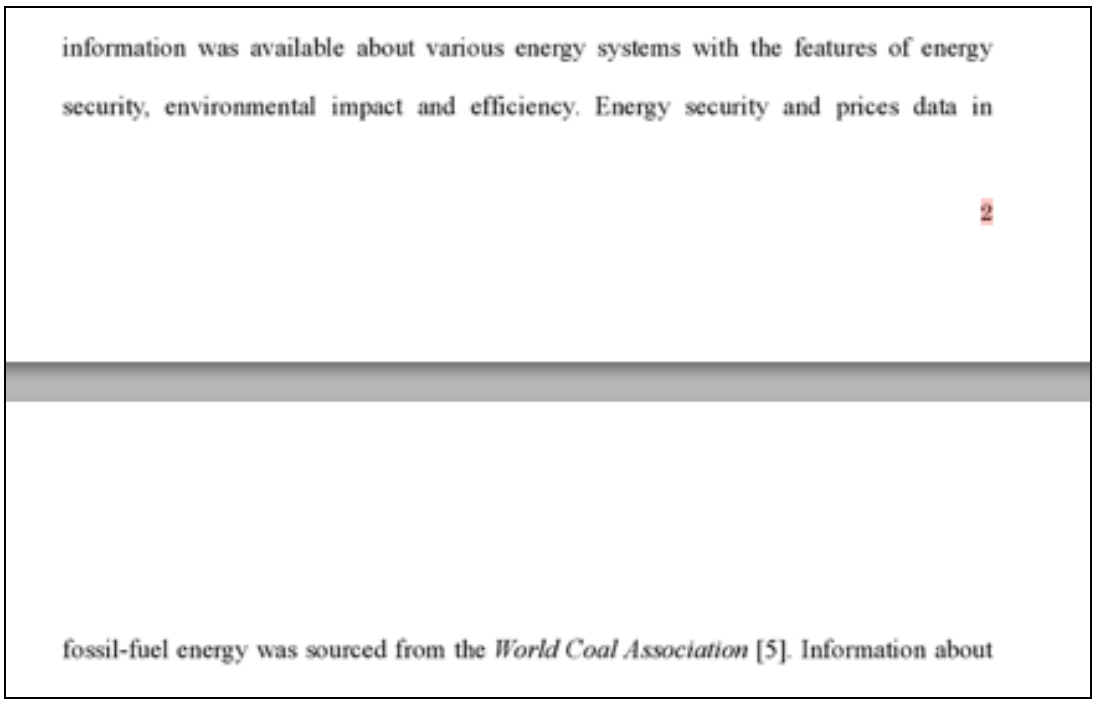


Figure A4. A screenshot showing page numbers flagged by Turnitin.

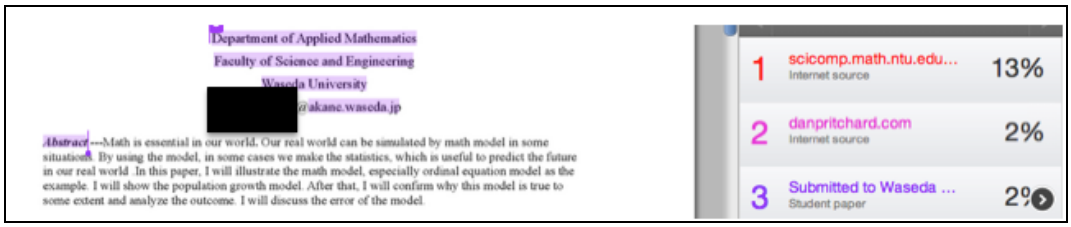


Figure A5. A screenshot of Turnitin showing a false positive for affiliation at 2%. The student's email address is blocked.

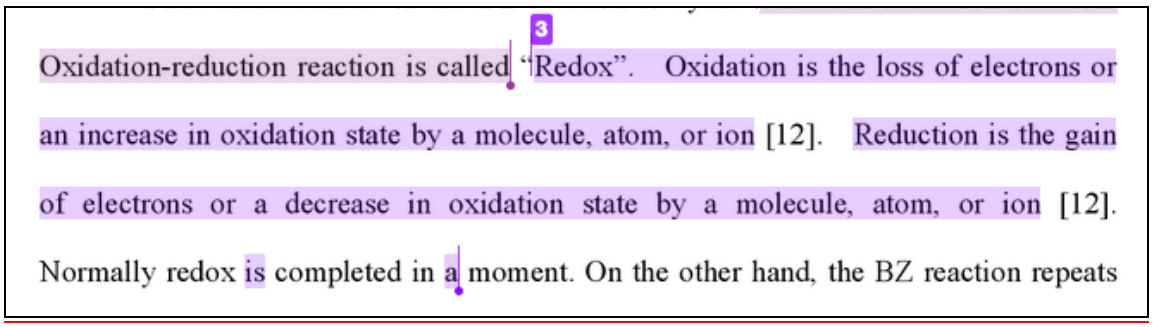


Figure A6. A screenshot of a general cited definition, flagged by Turnitin but allowed by the instructor.