Assessing the Accuracy of Plagiarism Detection Systems

Tim Vandenhoek
International Christian University, Japan

ABSTRACT

Plagiarism in academic writing is known to be an issue of concern for educators, administrators, and students alike. Using self-reporting studies and plagiarism detection software, previous research has established that plagiarism in university-level academic writing is relatively common amongst the work of L1 and particularly L2 writers and that paraphrasing and the various definitions of what constitutes plagiarism are notable challenges. However, previous research has been limited in the plagiarism detection systems used to evaluate student writing and has shed little light on their accuracy. This study seeks to contribute to scholarship on plagiarism by both measuring potential plagiarism in the writing of students at a university in Japan and assessing the accuracy of two plagiarism detection systems, Google Classroom and Grammarly. It was found that student writing was slightly less likely to feature content which may be deemed plagiarized, and the total amount of such content as a percentage of the total was notably low. The accuracy of the two systems differed greatly, with the Google Classroom version being more reliable in terms of generating fewer false positives and having more potential to be used as a self-editing or learning tool for students.

Keywords: *online learning; English for Academic Purposes; Plagiarism; Academic Writing*

BACKGROUND

The expansion of access to the Internet has led to greater concern about student plagiarism (Flowerdew & Li 2007). The Internet functions as something of a double-edged sword, both allowing for students to discover a wealth of sources which they can incorporate into their writing or, in some instances, copy from. It further gives administrators and instructors better means to detect any academic misconduct by use of plagiarism detection software (McKeever 2006; Youmans 2011). Given the uncertainty surrounding global education during the SARS COVID I9 pandemic, which at times saw more than 1 billion students out of the classroom (UNESCO 2021), the switch to online classes, and the continued growth of Internet access, concerns about plagiarism seem set to remain for the near future.

While plagiarism in academic writing is a well-researched topic, there is still a need to discover how accurate the systems used to detect plagiarism are and how they can be leveraged to better improve student writing and help learners avoid potentially damaging accusations of academic misconduct. The present study seeks to contribute to this discussion by evaluating the effectiveness of two plagiarism detection systems and their possible application as learning tools.

LITERATURE REVIEW

Plagiarism

Considerable research into plagiarism in L1 and L2 student writing has been conducted. One widely agreed upon notion is that defining plagiarism is troublesome, a topic on which several scholars have commented (Borg 2009; Bretag & Mahmud 2009; Flowerdew & Li 2007; Martin 2005; Merkel 2019; Pecorari 2015). This confusion exists not only within the relevant literature but also, unsurprisingly, between teachers and students (Keck 2014; Shang 2019). One reason for this lack

of agreement is that numerous behaviours deemed dishonest fall on a spectrum of plagiarism, including relatively minor issues such as poor paraphrasing to wholesale copying and contract cheating (Braumoeller & Gaines 2001; Howard 1995; Flowerdew & Li 2007; Nitterhouse 2003; Pecorari & Petric 2014; Walker 2005). This in turn leads to understandable frustration on the part of students (Flowerdew & Li 2007). In a finding that perhaps sums up the issue by leaving more questions than answers, an earlier study found that just over 40% of university staff respondents felt that they fully understood their own university's plagiarism policy, compared to 52% of students (Graham-Matheson & Star 2013).

One key idea in the literature is the necessity for a distinction between intentional and unintentional plagiarism in the ways in which cases of potential plagiarism are handled (Chien 2014; Flowerdew & Li 2007; Pecorari 2015; Shang 2019). This seems in part connected with the idea that acts of plagiarism are seen as an expected element in the work of developing writers, with it being variously characterized as a "survival strategy" (Flowerdew & Li 2007, p.168), a "healthy effort to gain membership in a new culture" (Howard 1995, p.236), evidence of "the desirable phenomenon of intertextuality" (Borg 2009, p.415), and "an integral stage for many students as they transition from novice to seasoned academic writers" (Merkel 2005, p.2). It has been suggested that plagiarism is more common in the work of less academically able L1 students (Keck 2014; Selwyn 2008), which lends credence to the idea that plagiarism is a developmental stage in writing. A slightly more cautious view is that dependence on or use of plagiarism should be viewed as stemming from a failure to learn the skills required for honest completion of a task and risks becoming cyclical: a student who does not have the ability to complete a task without plagiarising could be increasingly dependant on academic dishonesty in future assignments (Youmans 2011) and it is still noted as a serious issue worthy of censure in many universities.

Research has attempted to uncover the degree to which students plagiarize in their writing and found a wide range in the amount and type of student plagiarism. Studies in which students self-reported their own plagiarism show relatively high levels of academic dishonesty. In perhaps the largest study (over 70 000 students), McCabe (2005) found that 38% of undergraduate students admitted copying or rewriting sentences without citations, 7% reported reproducing full sentences without alterations or citations and 7% had submitted work written by someone else as their own. A survey of over 1200 L1 British university undergraduates (Selwyn 2008) indicated that over 60% admitted to some kind of online plagiarism in the previous year, with the most common breach being the unattributed copying of sentences from the Internet. Scanlon & Neumann (2002) surveyed nearly 700 undergraduate students in the United States, 24% of whom reported plagiarizing from online sources and 24.5% from printed sources. Smaller studies (fewer than 300 respondents) have noted similar rates of online copying for learners in Australia (32%, Sutherland-Smith 2008) and the UK (32.2%, Szabo & Underwood 2004). While noting that comparing data from different studies can be problematic, a rate of approximately 30% seems to be common, and several of these studies were rather dated. Given the expansion of the Internet in education, since the SARS COVID I9 pandemic and restricted access to libraries, it seems safe to assume that most plagiarism is now of online sources.

Actual detected levels of plagiarism (by use of plagiarism detection systems) in student writing are broadly like those of anonymous self-reporting, though again with a wide range of rates and some possible outliers. Analyzing written work (n = 129) from business graduate students using Turnitin.com, a commercial plagiarism detection service, Martin (2005) found some form of plagiarism in nearly half of assignments, with between 5-50% of the content of the paper matching another source. 26% of the 182 essays of mainly international students in master's degrees programs in research by Barrett & Malcolm (2006) had at least 15% matching based on alerts from Turnitin.com, a percentage equal to that found by Walker (2010) in a larger (n = 569) study of New Zealand undergraduates studying business, with roughly half being moderate (under 20% matched content) and half extensive (over 20% matched content) plagiarism. Data from two assignments (a

total of 125 papers) in Youmans (2011) showed 31% and 46% respectively, had 10% unoriginal content as determined by Turnitin.com. These findings, though, contrast with other studies that suggest lower levels of plagiarism. A more recent (2020) and far larger (n = 12 937) study of L2 students found rates of just 3.4%, though the author acknowledged that, due in part to the use of teacher intuition rather than a plagiarism detection system, the actual rates of plagiarism were likely higher (Perkins et al. 2020). Using a bespoke detection system, Warn (2007) found only 10.8% of written work (n= 74) was flagged as having between 3-15% plagiarized content.

As links have been seen between plagiarism and cultural values (Keck 2006), particularly those stemming from Anglo-Saxon concepts of intellectual property (Flowerdew & Li 2007), plagiarism is seen as a particular challenge for L2 learners, it having been "well established that plagiarism is relatively common in much L2 writing" (Pecorari 2015, p. 94). Studies show that the written work of L2 writers displays more instances of plagiarism. Walker (2010) found plagiarism in 37.5% of the work of international students, a rate twice as high as their native speaker counterparts. Shi's (2004) research similarly concluded that second language writers used twice as much (60%) textual borrowing in a summary writing task. Additionally, it has been noted that the kind of plagiarism in non-native writing is more direct, particularly the frequency of word-for-word copying (Keck 2006; Walker 2010). In a study of 34 EFL students, data from Shang (2019) discovered that 26% of content was plagiarized as determined by Turnitin. In a small study (n=32), Oda & Yamamoto (2008) found that over 90% of Japanese students submitted a written summary with over 40% similarity to the source text.

The problem of plagiarism does not immediately seem to be one of a lack of awareness or understanding of how seriously it is taken in universities; several studies have shown that both L2 teachers (Chien 2014; Hu & Sun 2016) and students (Balbay & Kilis 2019; Hu & Lei 2012; Rieber 2017; Wheeler 2009) can identify and disapprove of plagiarism in written work, though others have found some degree of tolerance towards plagiarism (Rinnert & Kobayashi 2005) and a possible knowledge deficit of plagiarism conventions (Keck 2006) in L2 students. Findings which suggest L2 writers understand and reject plagiarism represent either a methodological issue (few students, anonymously or otherwise, could be expected to express a tolerance for plagiarism, particularly in contexts in which they know it to be an institutional offence) or a problem of application. In a small study, university students reported understanding the rules of plagiarism but not how to apply them (Rieber 2017), a finding that deserves further investigation.

Paraphrasing

As academic writing is "largely reliant on the skill of paraphrasing to demonstrate that the author can capture the essence of what they have read, they understand what they have read and can use the appropriately acknowledged evidence in support of their responses", paraphrasing is seen as a vital tool (Rogerson & McCarthy 2017, p.2) and a challenge for L1 and L2 writers that requires a number of linguistic and content skills (Keck 2006; Pecorari 2015). One of the most challenging aspects of paraphrasing is the degree to which a source must be modified for it not to be considered plagiarism. Like notions of plagiarism itself, exactly what constitutes appropriate paraphrasing is the subject of much debate in the literature. Poor paraphrasing, that is, restating a source without making adequate modifications to its syntax or vocabulary, is generally considered to be a form of plagiarism, but, like the broader plagiarism debate, exists on a wide spectrum between clumsy or insufficient paraphrasing to something closer to outright plagiarism.

Thus, it is no surprise that paraphrasing is a notable challenge for students (Hu and Sun 2016; Oda & Yamamoto 2008; Shi 2004, Weigle and Parker 2012; Yamada 2003), and problematic paraphrasing is a feature of both L1 and L2 writing (Shi 2004; Walker 2010). Several factors may account for the difficulty faced by L2 writers when paraphrasing, the most obvious of which is general L2 proficiency. Successful paraphrasing depends on near-perfect understanding of the

source material and the ability to express that idea in new language. Therefore, a lack of competence may affect both their ability to paraphrase as well as their choice of paraphrase strategy, with L2 writers being seen as making fewer changes to source material incorporated into writing than L1 writers (Currie 1998; Keck 2006; Shi 2004), leading to potential accusations of plagiarism. Even advanced L2 writers struggled with accurately restating material, particularly in terms of expressing the author's intended epistemic stance (Basham & Rounds 1981). Instruction in paraphrasing in their home countries is seen to be lacking (Keck 2014), and the training they do receive may not put enough emphasis on the "inferential thought processes" that characterize advanced level paraphrasing (Yamada 2003, p.215).

The complex rules concerning citations and referencing in academic writing is another area in which L2 students are known to struggle (Chien 2014; Warn 2006; Weigle & Parker 2012) and, compounded by a lack of familiarity with or understanding of academic writing conventions in terms of appropriate paraphrasing (Currie 1998; Keck 2006), it should be expected that learners will struggle with paraphrasing. Combined with the fact that a failure to paraphrase properly may be considered an academic offence, instructing students in the rules, techniques, and guidelines that can help them avoid such incidents should be viewed as important, and online plagiarism detection systems may play a role in that process.

Plagiarism Detection Software

Just as the expansion of computing and the Internet has perhaps provided students with more means to plagiarise, plagiarism detection systems, have given instructors and administrators increased measures to detect any academic dishonesty. At their most basic, these services compare submitted student text to a database and attempt to find matching sections. Though teacher intuition and simple web searches for strings of suspicious text are also potential approaches (McKeever 2006), paid systems seem to be the preferred method, with Turnitin the most common service referred to in published research. It is a paid service that compares student submissions against open websites, open access and subscription academic databases and previously submitted student papers (Turnitin.com 2021).

Students (Balbay & Kilis 2019; Betts, Elder and Trueman 2012; Evans 2004; Grahame-Matheson & Star 2013; Rolfe 2011) and teachers (Chien 2014; Grahame-Matheson & Star 2013; Rolfe 2011) alike generally reported positive feelings towards plagiarism detection systems, both in terms of ease of use and effectiveness in identifying or dissuading plagiarism. It has been noted that such software can be useful for students in stopping themselves from unintentional plagiarism (Ledwith & Risquez 2008; McKeever 2006; Shang 2019).

Concerns, though, have been raised regarding various aspects of these systems, including their accuracy (Heckler, Rice & Hobson 2013), and the body of texts that they compare student work against (McKeever 2006). It is also accepted that plagiarism detection systems alone are not sufficient, and that teacher assessment of alerts is needed due to the occurrence of false positives (Betts et al. 2012; Bretag & Mahmud 2009; Graham-Matheson & Star 2013; McKeever 2006; Perkins et al. 2020; Warn 2006). For example, Barrett & Malcolm (2006) found that over one-third of the alerts produced by Turnitin could not be categorized as genuine plagiarism due to the system not recognizing quotation marks, references (presumably in the works cited section of the document) being flagged as problematic or other reasons). Few other studies address the issue of accuracy, though in all systems there remains the risk of both false negatives and positives (Perkins et al. 2020).

Several studies (each using different methodologies and, crucially, definitions of plagiarism) have been conducted attempting to assess the efficacy of using such systems to reduce academic misconduct by students. Studies of L1 writers by Martin (2005), Betts et al. (2012) and Heckler et

al. (2013) showed that a program of training and use of plagiarism detection software led to lower levels of detected plagiarism. Shang (2019) recorded similar results for L2 learners, though failed to show a connection between awareness of plagiarism conventions and decreased occurrence levels and participants were told that plagiarism could lead to a grade of zero on the assignment, which may have affected results. In a larger, longitudinal study of L2 learners, Perkins et al. (2020) found a 37% reduction in all forms of plagiarism after training and intervention, though that study relied on teacher intuition rather than software to detect suspected cases of academic misconduct. Walker (2010) found less plagiarism in the work of both L1 and L2 students after the introduction of plagiarism detection software but noted that more than 20% of students displayed the same behaviour even after being made aware of the use of the system and the fact that plagiarism was found in their first assignments. In a contrasting finding, Youmans (2011) discovered no change in student plagiarism after learners were made aware of the introduction of detection systems.

Use of such systems, however, is not seen as limited to instructors finding problematic academic behaviour but also viewed as a teaching tool that can help students avoid such issues. Several authors have advocated the use of this software to help students remedy unintentional plagiarism (Shang 2019) or to deal with it in a non-disciplinary, remedial teaching-based approach (McKeever 2006) while noting that students require training in revising their work based on results returned by detection systems, particularly in what constitutes a false positive and can be safely ignored (Rieber 2017).

Summary of Research and Rationale for Present Study

Based on the research surveyed above, several points seem salient. First, plagiarism is a potential issue, particularly for L2 writers, specifically connected to paraphrasing. Further, there is substantial, though not conclusive, evidence which suggests that plagiarism detection software can be useful for identifying potential cases of misconduct and to help students avoid accusations of plagiarism stemming from poor understanding of the conventions of paraphrasing, using sources in academic writing and proper use of sources with citations. In a review article, Pecorari (2015) called into question the need for further broad research into levels of plagiarism in student writing but highlighted the need for studies which can help students avoid plagiarism and develop their writing skills. There are no studies known to the author that have assessed the effectiveness of the plagiarism detection systems offered by Google Classroom or Grammarly.

The present study seeks to further that goal by evaluating the tools available to a group of students, namely the plagiarism detection systems which are part of Google Classroom, a popular learning management system (LMS), and Grammarly, a common online grammar and spelling check system which features a plagiarism check in its paid version. While much research has been conducted into the rates of plagiarism in student writing, little light has been shed on the accuracy of the systems used to detect it, and most studies have been predicated on the use of Turnitin. While it does seem something of a global institutional standard, it is a paid system, which limits access. The effectiveness of the plagiarism detection system integrated into Google Classroom, a free service whose use has seen a dramatic increase during the COVID 19 pandemic (De Vynck & Bergen 2020), has seemingly not been assessed by scholarly research. To direct students to resources which could help them, those resources must first be properly assessed to better understand their functions and possible shortcomings.

METHODOLOGY

Setting

This study takes place in a small, bilingual, private liberal arts university in Tokyo, Japan. The sample used in this study is from learners in Stream 3, which is the largest grouping, including students who have TOEFL ITP scores between 450 and 580 or an IELTS band score of 4.0-5.5. In their first year of university, students in this stream take three semesters of English for Academic Purposes classes, having seven periods of 70 minutes per week.

The written work for this study was digitally submitted in the last part of the final semester and is the third source writing assignment in the course. The writing used in this study was an argumentative essay concerning an issue related to bioethics. The assignment specified a word limit of 1000-1200 words using at least eight English-language sources. The assignment included an outline, a first draft and a final draft. Students were asked to allow the first draft of their essays to be analyzed in the study, and no incitements or compensation was offered. Student co-operation in the research process was high, with 28 of 30 learners agreeing to take part. Twenty-seven essays were analyzed for this research (one student did not complete the task, and that essay could not be used in the study). Two others did not complete a first draft and their final draft was used in the study instead. This relatively modest-sized sample allows for detailed examination of each instance of possible plagiarism highlighted by the two systems, thereby making it well-suited to assessing the efficacy of the detection software.

Plagiarism Detection Systems

Two plagiarism detection systems, Google Classroom Originality Reports and Grammarly, were used to analyze the student submissions. In Google Classroom, a free learning management system, the system is termed "Originality Reports" which can only be used for five assignments per class as part of the free model, though under a paid model, there is no limit. The system is optional, and the teacher must turn it on when creating an assignment. Students are also able to perform and view an originality report before submitting an assignment. As all assignments in the course are submitted via Google Classroom, use of the plagiarism checker does not require many extra steps for students or teachers. The system checks submissions against "all pages accessible by Google Search and Google Books" (Classroom Help 2021) and thus does not include any journal articles behind paywalls or requiring subscriptions. Unlike Turnitin, the system does not keep a database of previously submitted student tasks and does not check for student-student collusion. The reports contain a total number of matches (for the purpose of this study, "match" will be used to define any alert of potential plagiarism linked to a possible source text generated by a plagiarism detection tool) and the total percentage of matched content. Matches are divided between "Flagged", content matching web pages and "Cited/Quoted", sections matching documents used by students as sources in the assignment (Classroom Help 2021). There is a toggle switch which allows Cited/Quoted matches to be ignored. In Google Classroom, the report highlights the matches and provides links to the web sources that match the text. This report expires after 45 days, though a PDF version can be printed or saved.

The second plagiarism detection system used was from Grammarly. All students and staff at the university have access to the paid version of the system which includes a plagiarism check. The system requires students to submit their text to the Grammarly webpage, which they are encouraged to do to check their spelling and grammar, but the plagiarism check is not automatic, and students must click on the button to perform the check.

Grammarly does not categorize the matches in any way nor compare the work to any other previously submitted student writing. There is no limit to the number of checks that can be

performed. The premium service is said to compare submitted text to "billions of web pages and academic papers in private databases" (Grammarly Blog 2020). Information from the free version states that this includes the ProQuest academic database (Grammarly 2021). Like Google Classroom, matches are highlighted in the text and links to matching content appear. The system shows the percentage of matched content and a list of possible matching sources.

Analysis

As stated above, plagiarism detection systems produce false positives and teacher intuition is needed in evaluating matches. Thus, each match was individually assessed and broadly categorized as no concern (false positive) or potentially problematic. Exactly what constitutes plagiarism is, as noted above, a question of some debate within the literature and previous studies have used vastly different standards for identifying potential misconduct. Some studies considered one unique word occurring in both the source text and assignment to be a concern (Keck 2006), two content words (Shi 2004) or three items in general (Rieber 2017) as the criteria while others (Rolfe 2011) considered anything below four lines of direct copying to be acceptable. As student participants in this study were most recently taught that proper paraphrasing should include a maximum of three words in the same order as the source text, any student writing in which at least three words occur in the same order as a source text that is not part of a direct quotation was subject to further analysis. This requires the somewhat subjective intuition of the teacher to assess the uniqueness of the phrase, its presumed level of language proficiency and the source of the match. As will be seen, several matches, particularly from Grammarly, are of common strings of words (chunks) or come from source texts that have no connection to the essay topic and therefore do not arouse any suspicions of plagiarism.

Thus, only any match in student writing which:

        a) is not part of an attempted direct quotation

and

        b) contains at least three identical words occurring in the same order as a
        source text

and

        c) which contains particularly unique strings of language connected to the topic of the
        essay

and

        d) is clearly above the perceived linguistic level of the student

or

        e) is from a source which generates some level of suspicion

will be considered potentially problematic.

These definitions are slightly ad hoc but given the overall goal of this research (assessing the systems, not the students or their intentions), seem defensible. As the goal of this study is to evaluate the two systems and their potential applications in a process writing environment, no attempt will be made to definitively assess any instance of student writing as plagiarism, as such a determination requires thorough knowledge of student intentions. "Potentially problematic" does not imply plagiarism; rather, these are sections which the students would need to modify prior to submission of the final draft.

The following research questions form the basis of this study:

Research Question #1   How common is potential plagiarism (matches) in student
                                    writing?

Research Question #2	What is the accuracy of each system- what percentage of the matches after further analysis are potentially problematic, and what kinds of false positives are generated by each system?
Research Question #3	What is the possible utility of each system as a teaching or self-editing tool for students?

RESULTS

Total Matches and Matched Content Percentage

Table 1: *Matches from Google Classroom Originality Reports*

Paper #	Flagged Matches	Quoted Matches	Cited Matches	Total Matches	Matched Content Percentage
4	1	0	0	1	2
6	4	0	2	6	7
7	1	0	2	3	3.8
10	1	0	1	2	3
12	5	3	8	16	38
21	1	0	0	1	1.8
24	1	0	0	1	2
26	1	2	0	3	4
Total	15	5	13	33	
Average per paper (n=27)				**1.2**	**2.3%**

While the number of papers that had matches and the total number of matches were both moderate in the Google Classroom Originality Reports, the amount of matching content as a percentage of the papers themselves was very low. As seen in Table 1, Google Classroom Originality Reports produced alerts for eight of 27 papers (30%). A total of 33 matches were found, split between "Flagged" (15) and "Quoted" (3) and "Cited" (15). There was an average of 1.2 matches and 2.3% matched content per essay in the sample, though this figure is heavily skewed by Paper 12, the removal of which reduces average matches per submission to 0.7 and matched content to 0.9%.

Three papers had one alert each, and only two had over three, both of which seem to be outliers in this collection of writing (Paper 12 with 19 and Paper 6 with 6). The same papers were the only ones identified as potentially having more than 5% content matching and, interestingly, were the two final drafts that were analyzed because no first draft was submitted.

The Grammarly system (Table 2) returned a higher number of total matches and papers with matches, though a similar percentage of matched content. The check returned alerts on twenty papers, 74% of the total. Grammarly results show an average of 1.7 matches per paper and an average matched content figure of 2.2% (removing Paper 12 from the calculations reduces those to 1.4 and 1.2%, respectively). Nine papers had only one match, and only five papers had over three. Only two essays had over 5% of potentially unoriginal content.

Both systems identified the same paper (Paper 12) as being the most problematic in terms of numbers of matches (Google-18; Grammarly-9) and content percent (Google- 38%; Grammarly-27%). The largest discrepancy between the systems was found in Paper 24 (one match, one percent in Google, seven matches, seven percent in Grammarly).

Table 2: Matches from Grammarly Plagiarism Check

Paper Number	Matches	Matched Content Percentage	Paper Number	Matches	Matched Content Percentage
2	1	1	16	1	1
3	1	1	17	4	3
4	1	1	18	4	4
6	3	3	19	2	1
7	1	1	22	3	3
8	1	1	23	2	3
12	9	27	24	7	7
13	1	1	25	1	1
14	1	1	27	1	1
15	2	1			
			Total	45	
Average per paper (n=27)				1.7	2.2%

Accuracy and False Positives

The data in Table 3 shows that the overall accuracy of the Google Classroom system was 58%; that is, of the matches produced, after further analysis, 58% were deemed to be potentially problematic. This lowers the total average number of matches per paper from 1.2 to 0.7, with matches in 22% of submissions. The accuracy of the "Flagged" matches was highest at 73%, while that of the "Quoted" was lowest (0%) though there were only three matches in that section.

Table 3: Google Classroom Originality Reports Accuracy

	Matches	False Positives	Potentially Problematic
Flagged	15	4	11
Average	45%	27%	73%
Cited	15	7	8
Average	45%	47%	53%

Quoted	3	3	0
Average	10%	100%	0%
Total	**33**	**14**	**19**
Average		**42%**	**58%**

Most false positives (see Table 4) identified by the Google Classroom system (six of 14) seem to be the result of unclear or irregular use of quotation marks, particularly in cases where students use direct quotes which themselves include direct quotes. The following example, for instance, contains only three quotation marks, which could lead to a false positive.

> *"What we found is that instead of the mutation being fixed, the chromosome carrying the mutation is gone" — a profound change that likely dooms the embryo, Egli said. Many other cells showed changes in other chromosomes that also could do harm."* (Lab Tests Show) (Paper #12)

Table 4: *Google Classroom Originality Reports False Positives*

Type	Amount
Chunk	1
Same words in a different order	1
Not in works cited	1
Indirect quote from textbook	2
Quote from textbook	1
Embedded quote	6
Punctuation/Spacing	2

In another case, the student used four double quotation marks rather than single quotation marks to identify the embedded quote:

> *According to Hercher, "it would be expensive- costs for IVF in the US average over $20,000 for each try, and testing can add $10,000 or more. And it would require an unpleasant two-week process of ovarian stimulation and egg harvesting. "It wasn't the way I saw myself making a baby," Olivia told me. But they wanted what the procedure could offer them: a guarantee that dystonia was eliminated for the next generation, and beyond."* (Hercher) (Paper #12)

It is possible that the system thus identified this as two separate quotes with plagiarised text between them.

Several false positives (three) are the result of the course textbook, which is composed of reprinted excerpts from other books and has no online analogue. Two false positives came from the use of indirect quotes. In two papers, students correctly quoted sections of the course textbook that quoted other sources (the philosophers Aristotle and John Stuart Mill). The Google Classroom software thus matched this text to the original source text or reproductions of the quotes found on various

websites (in one case, the social media site Pinterest) and produced an alert. The other case was a correctly cited quotation that the system matched to the text that was reprinted in the coursebook.

In one other case, the system identified a source text using the same words, but in a different order (emphasis as in Originality Report):

Student writing:	*For instance**, the** EU regulation requires labeling all **food** products **containing**, produced from, or containing ingredients from GMOs (Paper #7)*
Source text:	*The Japanese authorities have required designated agricultural Products and processed food items containing **GM materials** to be labelled. **For the** processed **food** items, those ingredients **containing** GM.*

Another match seems to be a case of coincidence in which the student and a text not used in the essay paraphrased the results of the same study in a similar way (emphasis as in Originality Report):

Student writing:	***the fertility rate has** been dramatically dropping, and according to the Institute for Health Metrics and Evaluation at the University of Washington, the live births per woman, which was **4.7 in 1950**... (Paper #24)*
Source text:	*... by the University of Washington's Institute for Health Metrics and Evaluation. ... The average number of children that a woman gives birth to - **the fertility rate – has** fallen from **4.7 in 1950** to...*

The sections that led to the match are lexical chunks ("the fertility rate") and the numerical data finding and are not seen as problematic.

The remaining false positives were the result of poor use of punctuation or spacing (two cases, one using a colon rather than quotation marks to introduce a direct quotation) and a quote properly referenced in the main text of the essay but without an accompanying entry in the works cited section (one).

Accuracy- Grammarly

The overall accuracy of the Grammarly system, seen in Table 5, was 20%, which means 22% of papers had potentially problematic sections, an average of 0.3 matches per essay. As the false positives produced by Grammarly differed vastly from those of the Google Classroom check, they were further analyzed by type of match. Three categories were found, a "Chunk" - a string of words which commonly appear as a group, a "Topic Chunk" - a chunk including specific reference to the topic of the essay, and "Quoted" - matches that occur within quotation marks.

Table 5: Grammarly Plagiarism Checker Accuracy

	False Positives			Potentially Problematic
	Chunk	Topic Chunk	Quoted	
Number	22	10	4	9

Percent	49%	22%	9%	20%

The most common type of false positive from the Grammarly results (49%) is a relatively long lexical chunk consisting of between five and ten words with no direct connection to the essay topic matching an unrelated source text. As seen in Table 6, it would be difficult to judge any of the phrases as especially unique; indeed, they are the kind of phrases one would expect to find in academic writing and, as the level of language does not raise any suspicions, and the identified source texts have no reasonable connection to the essay topic, they are of no concern.

Table 6: *Grammarly Plagiarism Checker False Positives- Chunks*

Example	Source text
There have been an increasing number of issues ... (Paper #2)	Recipe
According to the survey conducted by Pew Research Center (Paper #18)	Law blog

In the second set of examples shown in Table 7, while the phrase does have a clear link to the topic of the essay (animal rights), it is the kind of language which may commonly occur in essays on the topic and thus considered to be a "Topic Chunk".

Table 7: *Grammarly Plagiarism Checker False Positives- Topic Chunks*

Example	Source Text
Animals also have rights like human beings, and... (Paper #16)	News site article with no connection to essay topic
An organization called People for the Ethical Treatment of Animals (Paper #18)	Essay Mill

Similarly, the proficiency of the language is again at the level of a high intermediate student and thus is categorized as no concern. This kind of match accounts for 22% of false positives. While the second example of this type was matched to a concerning source (an essay mill website introducing several essays on various themes connected to animal rights), it would seem prejudicial to accuse a student of plagiarizing only one wholly non-unique phrase from such a source.

In contrast, in the third set of examples seen in Table 8 below, the strings of text, while relatively uncomplex, are longer (10 words and over) and match suspicious sources (a full essay on an essay mill website and a full essay written by an undergraduate student, both of which are directly connected to the essay topics) and are thereby deemed potentially problematic. However, this represents a low bar for plagiarism, and both could be considered borderline cases and, as it is essentially impossible to ascertain whether a student could have accessed a particular website, in practice, both students should be accorded the benefit of the doubt.

Grammarly is slightly more successful in handling quoted content than Google Classroom, with only 4 false positives in that category. All four also produced alerts in Google Classroom for the same reasons (issues with quotation marks and properly cited sections of the course textbook), the only overlap between the two systems in terms of false positives.

Table 8: Grammarly Plagiarism Checker Potentially Problematic Matches

Example	Source Text
On the other hand, in countries where abortion is legal... (Paper #6)	Essay Mill
There are two kinds of euthanasia, active and passive. Active euthanasia is… (Paper #13)	University website which shares undergraduate student research

False Negatives

By their nature, it is impossible to identify all false negatives, nor is it the goal of this research. However, it is important to recognize that instances of potentially problematic writing can go undetected by the two systems. The following are some of the four false negatives identified by the instructor.

Paper 6: *Studies have shown that pregnancy is often followed by vomiting, gestational diabetes, hemorrhoids, bowel problems, incontinence and 95% of first-time mothers experience vaginal tearing (Wright).*

Source: *...the gestational diabetes, hemorrhoids, bowel problems, incontinence, or any of the common complications that follow pregnancy and birth. Even if you are blessed with an easy pregnancy, some reports say as many as 95 percent of first-time mothers experience vaginal tearing.*

In this case, there is notable overlap between the student essay and the source text, which is correctly cited, though neither Grammarly nor Google Classroom identified as a potential match though the content itself, an online magazine article, is not behind a paywall, and perhaps shows the difficulty of paraphrasing medical or technical terms.

Other examples show a similar issue of poor paraphrasing or quotation practices:

Paper 7: *In particular, Vitamin A deficiency causes up to 500 000 cases of childhood blindness and 2-3 million deaths annually (Hug 89).*

Source: *Vitamin A deficiency is a serious burden on the health of millions of children living in developing countries who cannot afford alternative sources of the vitamin, and it causes up to 500 000 cases of childhood blindness and 2–3 million deaths annually.*

Paper 7: *Some studies have revealed toxic effects or signs of toxicity (Hilbeck et al. 2).*

Source: *Independent studies of this type are rare, but when such studies have been performed, some have revealed toxic effects or signs of toxicity in the GM-fed animals.*

These could be thought of as the type of issue plagiarism systems should be able to detect, particularly as the source text in both instances are open-access journals which are in the works cited list of the essay. These false negatives show no system can be reliably expected to highlight all potential instances of plagiarism within writing and, as with assessing the matches that the systems do produce, that teacher intuition is indispensable.

DISCUSSION

Research Question #1- How common is potential plagiarism (matches) in student writing?

Data from this study suggests that while plagiarism within student writing occurs at a similar frequency found in previous studies, the type and amount of matched content are very low. Comparing results across studies with different methodologies is fraught with difficulty, though rates of potential plagiarism in this corpus (22% of papers, as calculated after secondary analysis) are like those discovered in studies of self-reported plagiarism or by use of plagiarism detection software. In the studies surveyed above, self-reported rates of plagiarism range from 24% to over 60%, with several studies finding that approximately 30% of students admit to plagiarism of some form. Detected levels of plagiarism in student writing range between 25% and 50%. With the caveat that the plagiarism checks performed in this study do not have the potential to identify student-student copying, it seems that plagiarism is less of an issue within this group of students.

Where the findings of this study diverge from previous research is in the amount of content in each paper that matches another source. After adjusting for the number of false positives, it is estimated that between 0.4% and 1.3% of total content in the essays corresponds to other sources (as both systems produce a matched content percentage for each paper in total rather than by individual match, these are only rough estimates), though it should again be noted that these figures are heavily influenced by one paper which seems to be an outlier in this data set. Removal of that paper from the data reduces the rate to between 0.3% and 0.5%. In sharp contrast, existing research shows a range of between 3-50%. Even accounting for the inability of Google Classroom or Grammarly to identify possible peer-to-peer plagiarism, this result seems notable.

It should again be noted here that the description of "potentially problematic" does not mean that these matches are evidence of plagiarism and in fact, several of those cases are borderline issues which would, in an actual teaching situation, may not require revisions before inclusion in a final draft. None of the matches found in this study would lead to any academic disciplinary procedures. In fact, 10 of the 19 potentially problematic matches identified by the Google Classroom check have either a parenthetical citation or some kind of reference (oblique or otherwise) to the source. This lends further weight to the argument that these matches represent unintentional plagiarism or inadequate paraphrasing.

Research Question #2- What is the accuracy of each system- after further analysis, what percentage of the matches are potentially problematic, and what kinds of false positives are generated by each system?

The two systems had notably different levels of accuracy. The integrated system which is a part of the Google Classroom assignment function was 58% accurate based on the somewhat strict conditions established in the methodology section, while the same figure for Grammarly was 20%. The results from Google can be further analyzed by type of match. Matches under the "Flagged"

category were the most numerous (15) and the most accurate (73%). "Cited" and "Quoted" matches were 47% and 0% accurate, respectively. Few other studies on the accuracy of plagiarism detection systems exist, but Barrett & Malcolm (2006) found that one-third of matches from Turnitin were false positives and highlighted the inability of the system to identify quotation marks consistently and correctly.

In assessing the kinds of language that result in false positives, the Google Classroom system returned a high number (over 40%) of false positives for quotations embedded with other quotations. As these were all within one paper, it could potentially be a minor issue that substantially affected the results of this study (removing them increases accuracy to 70%). Indirect quotations and other issues specific to the coursebook, errors in punctuation or spacing were the other most common false positives. Crucially, the types of false positives that arise from the Google Classroom system appear to follow patterns and based on simple guidelines, could be somewhat reliably assessed, and ignored by students (see below).

Conversely, more than 70% of the Grammarly false positives would require some level of intuition to confidently identify as they deal with lexical chunks, and it would be challenging for a student to evaluate whether these were sufficiently unique to warrant changes or the use of quotation marks.

Research Question #3- What is the possible utility of each system as a teaching or self-editing tool for students?

Given its relatively high accuracy and the types of false positives it generates, the Google Classroom system has clear potential as a teaching and learning tool in helping students avoid unintentional plagiarism prior to submission of their written work. Provided with a relatively short and simple set of instructions, it could be possible for students to produce an originality report for their work, assess which of the matches are false positives and can be ignored and which sections of their text require revisions. The overall message to students should be that the system itself is a tool and not entirely reliable and that, along with false positives, can feature false negatives and that the best way to avoid potentially problematic sections within their submissions is a cautious approach to the use of sources, paraphrasing and citations.

Ideally, the students would submit their originality reports to the teacher before the lesson to allow the teacher to identify any common issues (with either false positives or problematic sections) and prepare any remedial or additional teaching materials. For example, in the case of the sample used in this study, there would be a need to review points such as correct use of quotation marks for embedded quotes and how to cite a source indirectly. Students could also be taught any rules of thumb, in this case, that they can safely ignore matches in which they quote from the textbook, but a match is found for the original source.

Depending on class level, the teacher could prepare a worksheet with several matches or opt for a freer practice of students assessing their own originality reports. In either case, the class should start with an explanation of how the system works, a review of the general rules for paraphrasing or citing sources and information on false positives. The students should be presented with several instructive examples of false positives and problematic sections, which would then be discussed and corrected as a group. The lesson would then proceed to group work and could finish with two groups joining together, showing either the matches they deemed false positives or problematic sections that they then rewrote. The feedback section of the class should give students the opportunity to ask for the teacher's judgement on any borderline or unclear cases. If they had not already done so, learners would then be encouraged to make changes to their own essays and run another originality report. Google Classroom allows students to generate five reports per assignment, so this could be an iterative process designed at making the submission as free from potential plagiarism as possible.

It is, though, more difficult to argue that recommending or requiring students to use the Grammarly system would be an appropriate and effective policy as training students to be able to reliably assess the Grammarly matches would require a significant investment of time and effort not justified by the low accuracy of the system. First, the overall accuracy of Grammarly is so low that there is a strong efficiency argument against its use. Of the nine potentially problematic sections of text identified by Grammarly, four were also found by the Google Classroom check. Of the remaining matches, three were considered potentially problematic by the strict standards of the methodology of this study, though are in practice borderline cases which may not require change or intervention. In total, then, the Grammarly check only resulted in two unique (that is, not noted by the Google Classroom system) potentially problematic matches from the overall total of 45.

The second reason against requiring students to submit their writing to the Grammarly system is the type of false positives it produces. As noted above, judging these matches to be of no concern requires far more intuition than doing the same for the false positives from the Google check. Novice writers cannot, and cannot be expected to, know what constitutes unique strings of language or which sources may appear to be suspicious to an instructor and, given the difficulty of source writing, the addition of any potential confusion or frustration could be counterproductive.

Limitations

These results are by their nature limited by the effectiveness of the systems being analysed; there is no perfect system that they can be compared against. The findings are further limited by the existence of false negatives (incidences of potential plagiarism which are detected by the researcher but not by either system). There is a possibility that some false negatives were not identified. As indicated by the participation rate, the participants of this study were in general cooperative and, due to the role of the instructor in grading and reading various stages of the assignment, may represent a group less likely to attempt any acts of academic dishonesty. Therefore, the results may not be suitable for comparison with other groups of learners.

CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

While student plagiarism is an issue, enough is now known about what forms that misconduct can take, how frequently it occurs and, to a lesser extent, why it exists in student writing. Pecorari's argument that more research is needed into ways in which students can be taught to effectively avoid accusations of plagiarism is well-taken, and with that in mind, there are several areas of inquiry which seem under-represented in the relevant literature. First, there is a limited amount of scholarship attempting to assess the accuracy of all plagiarism detection systems, even of Turnitin, by far the most researched software. For students to take advantage of these systems, their strengths and weaknesses must be better understood, particularly the kinds of false positives that are likely to occur. Future research should focus more on the closer examination of the passages flagged by these systems and whether they could be considered instances of potential plagiarism. As use of Google Classroom is expanding, further investigation into its plagiarism detection system, particularly how students can leverage these tools in order to improve their own work before submission, would be of benefit to teachers and students alike. Finally, like studies performed on the use of Turnitin, the student experiences of using other types of plagiarism detection software are another area in which further research could be illuminating.

REFERENCES

Balbay, S. & Kilis, S. 2019, "Perceived Effectiveness of Turnitin in Detecting Plagiarism in Presentation Slides", *Contemporary Educational Technology,* vol. 10, no. 1, pp. 25.

Barrett, R. & Malcolm, J. 2006, "Embedding plagiarism education in the assessment process", *International Journal for Educational Integrity,* vol. 2, no. 1, pp. 38-45.

Basham, C. & Rounds, P. 1984, "A discourse analysis approach to summary writing", *TESOL Quarterly* vol. 18, no. 3, pp. 527.

Betts, L., Bostock, S.J., Elder, T.J. & Trueman, M. 2012, "Encouraging good writing practice in first-year psychology students: An intervention using Turnitin", *Psychology Teaching Review,* vol. 18, pp. 74-81.

Borg, E. 2009, "Local plagiarisms", *Assessment & Evaluation in Higher Education,* vol. 34, no. 4, pp. 415-426.

Braumoeller, B.F. & Gaines, B.J. 2001, "Actions Do Speak Louder than Words: Deterring Plagiarism with the Use of Plagiarism-Detection Software", *Political Science and Politics,* vol. 34, no. 4, pp. 835-839.

Bretag, T. & Mahmud, S. 2009, "A model for determining student plagiarism: Electronic detection and academic judgement", *Journal of University Teaching and Learning Practice*, vol. 6, no, 1, pp. 49-60.

Chien, S. 2014, "Cultural Constructions of Plagiarism in Student Writing: Teachers' Perceptions and Responses", *Research in the Teaching of English,* vol. 49, no. 2, pp. 120-140.

Currie, P. 1998, "Staying out of trouble: Apparent plagiarism and academic survival", *Journal of Second Language Writing,* vol. 7, no. 1, pp. 1-18.

De Vynck, G. & Bergen, M. 2020. "Google Classroom Users Doubled as Quarantines Spread". *BloombergQuint.* Available at: https://www.bloombergquint.com/business/google-widens-lead-in-education-market-as-students-rush-online. (Accessed August 2021).

Evans, R. 2006, "Evaluating an Electronic Plagiarism Detection Service", *Active Learning in Higher Education,* vol. 7, pp. 87-99.

Flowerdew, J. & Li, Y. 2007, "Plagiarism and Second Language Writing in an Electronic Age", *Annual Review of Applied Linguistics,* vol. 27, pp. 161-183.

Google Classroom Help. 2021. "How an originality report is created." Available at: https://support.google.com/edu/assignments/answer/9421894?hl=en. (Accessed August 2021).

Graham-Matheson, L. & Starr, S. 2013, "Is It Cheating or Learning the Craft of Writing?", *Research in Learning Technology,* vol. 21, no. 1, pp. 1-13.

Grammarly Blog. 2020. "How to check your work for plagiarism". Available at: https://www.grammarly.com/blog/plagiarism-checker/. (Accessed August 2021).
Grammarly. 2021. "Plagiarism checker by Grammarly". Available at: https://www.grammarly.com/plagiarism-checker. (Accessed August 2021).

Heckler, N.C., Rice, M. & Hobson Bryan, C. 2013, "Turnitin Systems", *Journal of Research on Technology in Education,* vol. 45, no. 3, pp. 229-248.

Hilbeck, A., Binimelis, R., Defarge, N., Steinbrecher, R., Székács, A., Wickson, F., Antoniou, M., Bereano, P.L., Clark, E.A., Hansen, M. and Novotny, E., 2015. No scientific consensus on GMO safety. *Environmental Sciences Euro*pe, vol. 27, no. 1, pp.1-6.

Howard, R.M. 1992, "A Plagiarism Pentimento", *Journal of Teaching Writing,* vol. 11, no. 2, pp. 233.

Hu, G. & Sun, X. 2016, "Chinese University EFL Teachers' Knowledge of and Stance on Plagiarism", *Comunicar,* vol. 24, no. 48.

Hu, G. & Lei, J. 2012, "Investigating Chinese University Students' Knowledge of and Attitudes Toward Plagiarism From an Integrated Perspective", *Language Learning,* vol. 62, no. 3, pp. 813-850.

Hug, K. 2008. Genetically modified organisms: do the benefits outweigh the risks? *Medicina*, vol. 44, no. 2, pp. 87–99.

ISAAA. 2019. "ISAAA brief 55-2019: Executive summary*"*. Available at: https://www.isaaa.org/resources/publications/briefs/55/executivesummary/default.asp. (Accessed August 2021).

Keck, C. 2014, "Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices", *Journal of Second Language Writing,* vol. 25, pp. 4-22.

Keck, C. 2006, "The use of paraphrase in summary writing: A comparison of L1 and L2 writers", *Journal of Second Language Writing,* vol. 15, no. 4, pp. 261-278.

Ledwith, A. & Risquez, A. 2008, "Using anti-plagiarism software to promote academic honesty in the context of peer reviewed assignments", *Studies in Higher Education,* vol. 33, no. 4, pp. 371-384.

Lei, J. & Hu, G. 2015, "Chinese university EFL teachers' perceptions of plagiarism", *Higher Education,* vol. 70, pp. 551-565.

Martin, D.F. 2005, "Plagiarism and Technology: A Tool for Coping With Plagiarism", *Journal of Education for Business,* vol. 80, no. 3, pp. 149-152.

McCabe, D. 2005, "Cheating among college and university students: A North American perspective", *International Journal of Educational Integrity,* vol. 1.

McKeever, L. 2006, "Online plagiarism detection services—savior or scourge?", *Assessment & Evaluation in Higher Education,* vol. 31, no. 2, pp. 155-165.

Merkel, W. 2020, "A case study of undergraduate L2 writers' concerns with source-based writing and plagiarism", *TESOL Journal,* vol. 11, no. 3, pp. 1-16.

Nitterhouse, D. 2003, "Plagiarism – Not Just an Academic Problem", *Teaching Business Ethics,* vol. 7, no. 3, pp. 215-227.

Oda, S. & Yamamoto, Y. 2008, "Paraphrasing: an Essential Tool for EAP", *Language Research Bulletin,* vol. 22, pp.77-88.

Pecorari, D. 2015, "Plagiarism in second language writing: Is it time to close the case?", *Journal of Second Language Writing,* vol. 30, pp.94-99.

Pecorari, D. & Petric, B. 2014, "Plagiarism in second-language writing", *Language Teaching,* vol. 47, no. 3, pp. 269-302.

Perkins, M., Gezgin, U.B. & Roe, J. 2020, "Reducing plagiarism through academic misconduct education", *International Journal for Educational Integrity,* vol. 16, no. 1, pp. 3.

Rieber, C., 2017. Using electronic plagiarism tools to teach university students to avoid plagiarism. In *Proceedings of the Conference on Plagiarism across Europe and beyond 2017.*

Rinnert, C. & Kobayashi, H. 2005, "Borrowing Words and Ideas: Insights from Japanese L1 Writers", *Journal of Asian Pacific Communication,* vol. 15, no. 1, pp. 15-29.

Rogerson, A.M. & McCarthy, G. 2017, "Using Internet based paraphrasing tools: Original work, patchwriting or facilitated plagiarism?", *International Journal for Educational Integrity,* vol. 13, no. 1, pp. 1-15.

Rolfe, V. 2011, "Can Turnitin be used to provide instant formative feedback?", *British Journal of Educational Technology,* vol. 42, no. 4, pp. 701-710.

Scanlon, P.M. & Neumann, D.R. 2002, "Internet plagiarism among college students", *Journal of College Student Development,* vol. 43, no. 3, pp. 374-385.

Selwyn, N. 2008, "'Not necessarily a bad thing …': a study of online plagiarism amongst undergraduate students", *Assessment & Evaluation in Higher Education,* vol. 33, no. 5, pp. 465-479.

Shang, H. 2019, "An investigation of plagiarism software use and awareness training on English as a foreign language (EFL) students", *Journal of Computing in Higher Education,* vol. 31, no. 1, pp. 105-120.

Shi, L. 2004, "Textual Borrowing in Second-Language Writing", *Written Communication,* vol. 21, no. 2, pp. 171-200.

Sutherland-Smith, W., 2008. *Plagiarism, the Internet, and student learning: Improving academic integrity*. Routledge.

Szabo, A. & Underwood, J. 2004, "Cybercheats: Is information and communication technology fuelling academic dishonesty?", *Active Learning in Higher Education,* vol. 5, no. 2, pp. 180-199.

Turnitin.com. 2021. "The Turnitin difference". Available at: https://go.turnitin.com/l/45292/2019-05-09/bj81hq/45292/180815/Turnitin___Our_Web_Crawler.pdf?_ga=2.189900395.1763782 112.1628473572-1519955557.1626218484. (Accessed August 2021).

UNESCO. 2021, "Education: From disruption to recovery 2020" Available at: https://en.unesco.org/covid19/educationresponse (Accessed August 2021).

Walker, J. 2010, "Measuring plagiarism: Researching what students do, not what they say they do", *Studies in Higher Education,* vol. 35, no. 1, pp. 41-59.

Warn, J. 2006, "Plagiarism software: No magic bullet!", *Higher Education Research &,* vol. 25, pp. 195-208.

Weigle, S.C. & Parker, K. 2012, "Source text borrowing in an integrated reading/writing assessment", *Journal of Second Language Writing,* vol. 21, no. 2, pp. 118-133.

Wheeler, G. 2009, "Plagiarism in the Japanese universities: Truly a cultural matter?", *Journal of Second Language Writing,* vol. 18, no. 1, pp. 17-29.

Wright, J. (2018). "Abortion is not murder*". Harper's Bazaar*. Available at: https://www.harpersbazaar.com/culture/politics/a19748134/what-is-abortion/. (Accessed August 2021).

Yamada, K. 2003, "What prevents ESL/EFL writers from avoiding plagiarism?: Analyses of 10 North-American college websites", *System,* vol. 31, no. 2, pp. 247-258.

Youmans, R.J. 2011, "Does the adoption of plagiarism-detection software in higher education reduce plagiarism?", *Studies in Higher Education,* vol. 36, no. 7, pp. 749-761.

_____