

The Role of Anti-Plagiarism Software in Learning to Paraphrase Effectively

Marina Dodigovic (mdodigov@gmail.com), Xi'an-Jiaotong Liverpool University, China

Abstract

Plagiarism, or using someone else's words or ideas without acknowledgment, is known to appear in university student writing. Culture, education, attitudes and language proficiency are viewed as the major causes of ESL writer plagiarism. The use of electronic plagiarism detection tools such as Turnitin¹ and SafeAssignment seems to provide a remedy. While previous research (Walker, 2010) demonstrates that Turnitin can be an effective plagiarism deterrent, this paper claims that Turnitin can be additionally used as a learning aid which, paired with focused instruction, could help the students improve their quoting and paraphrasing skills. The claim is supported by an action research case study which tracks the plagiarism levels in a writing-intensive course. The conditions include the use of Turnitin to upload (1) the final assignment drafts only, (2) all assignment drafts, and (3) all drafts after explicit instruction in the use of Turnitin to improve paraphrasing skills. Condition 3 appears to be the most effective.

Keywords: plagiarism, plagiarism detection software, second language writing, citation, paraphrase

INTRODUCTION

English-medium tertiary education programs emphasise the role of writing, thus making themselves vulnerable to plagiarism. In particular text borrowing or appropriation has caused librarians (Amsberry, 2010; Lund, 2004), literacy instructors (Evans & Youmans, 2000) and ESL specialists (Lankamp, 2009; Shi, 2006) to give it increasing research attention, with the focus on possible reasons and remedies. Two factors seem to have contributed to this development: the advent of computational plagiarism detection tools and the emancipation of ESL writing research. However, there seems to be little overlap between the two in the available literature. To bridge the gap, this study combines the application of electronic plagiarism detection tools with the instructional context of second language writing. Based on ESL student writing in first year composition classes at an English-medium university in the Middle East, the study compares the effectiveness of three different anti-plagiarism software uses, namely as a: 1) detector, 2) deterrent, and 3) learning tool.

PLAGIARISM IN SECOND LANGUAGE WRITING

While the availability of sources on the Internet has been identified as the precipitating factor in the increase of student plagiarism in general (Erkaya, 2009; Reingold & Bartz, 2011; Shi 2006), culture, education, attitude and language proficiency have been viewed as the major causes of ESL writer plagiarism (Amsberry, 2010; Basharina, 2007; Erkaya, 2009). Confucian cultures are believed not to have cultivated as strong notions of text ownership as Western societies (Lund, 2004; Shi, 2006), seeing textual borrowing as a sign of respect toward the original. To educational influences, such as rote learning (Ballard & Clanchy,

1984) or learning by imitation, which are often associated with Confucian cultures, Hyland (2001) additionally counted the misplaced cultural sensitivity of ESL instructors, which prevents them from giving clear and meaningful feedback on textual borrowing. On the other hand, writing in a second language and trying to accommodate a new rhetorical practice could be the linguistic causes of ESL student patchwriting. Patchwriting refers to seamless and unacknowledged incorporation of fragments from an original into one's own text (Amsberry, 2010). Finally, students may find themselves unmotivated to write or discouraged by their instructor's negative attitude toward writing (Erkaya, 2009).

Solutions are expected from developing comprehensive definitions of plagiarism, library orientations, classroom instruction, reference services and library websites (Amsberry, 2010). In particular classroom instruction suggestions are relevant to ESL writing instructors. While Lankamp (2009) proposed that such instruction should include more than "the content of a style guide" (p. 1), Erkaya (2009) believed that the students should be specifically taught research skills, in addition to quoting, paraphrasing and blending, in a step-by-step approach, always under the watchful eye of the instructor. Shi (2006) went a step further by pleading for a "postmodern pedagogy of imitation and student authorship" (p. 264). This pedagogy would go beyond the "mechanics of citation" toward "inferential thinking" (p. 277) when referring to the ideas of others. An appropriate method for achieving this is by teaching the students "repetition, imitation, appropriation and manipulation" (Shi, 2006, p. 277) of the words of others. On another note, Evans & Youmans (2000) pointed out the importance of socially co-constructing the concept of plagiarism through discourse with instructors and peers. An example of such discourse co-created by students through blogging was described by Bloch (2007). In addition to engaging students in constructive discourse about plagiarism, instructors, according to Hyland (2001), need to develop a clear and unambiguous feedback style.

THE ROLE OF PLAGIARISM DETECTION SOFTWARE

Erkaya (2009) made a reference to plagiarism detection software, otherwise rare in research dealing with ESL writer text appropriation. Tools such as Turnitin or SafeAssignment, to name just a few, can identify the fragments of text borrowed from other sources. While concerns are voiced over the merely punitive uses of such software, an empirical study (Walker, 2010) demonstrated that Turnitin can be an effective plagiarism deterrent. However, since pedagogy seems to be the most effective means of plagiarism prevention, according to Erkaya (2009) such software should also be used to support learning. It can provide feedback which is both unambiguous and direct, without the frequent cultural oversensitivity associated with the human instructors (Hyland, 2001). In this capacity, it almost becomes a dialogue partner, assisting the student in the construction of the plagiarism concept.

In order to determine the role of plagiarism detection software, it is necessary to understand the mechanisms of both student plagiarism and the plagiarism detection software. Dick et al. (2003) subsumed instances of copying from sources, in or outside exams, collaboration (collusion) and deception under the term plagiarism. Anderson and Poole (1994) interestingly stated that "a narrow line often separates plagiarism from good scholarship" (p. 27). While good scholarship may use the same sources and to the same extent as a plagiarised text, the skill of giving those sources a separate voice is what sets good scholarship apart. A document

crafted at an English-medium university in the Middle East attempts to draw exactly that “narrow line” (p. 27) that Anderson and Poole (1994) referred to:

Plagiarism is attempting to present another person’s work as a student’s own. If a student copies another student’s paper, copies a paper from a textbook, or downloads a paper from the Internet and presents it as his/her own work, this is plagiarism. If a student takes a sentence or even part of a sentence from an outside source without putting quotation marks around it and citing the source, this is plagiarism. If a student takes ideas, facts, and opinions from an outside source without citing the source, this is plagiarism also. (AUS, 2007)

The above text goes beyond definition, trying to delimit the scope of plagiarism. Pursuing the same goal, another university manual goes a step further by adding self-plagiarism to the list.

...plagiarism is defined as: (i) extensive use of paraphrasing, quotations, etc; (ii) the partial or full use of previously published work or essays; (iii) representing the designs and concepts of another person without citation; (iv) the repeated use of small amounts of paraphrasing, quotation, etc. without citation. (XJTLU, 2012)

The latter also seems to reflect the confidence of the institution in its own ability to identify which essays have been previously published or submitted. Such confidence has been increasingly connected with the use of plagiarism detection software (Walker, 2010). This type of software compares the target text with a variety of electronic sources. These usually fall into two large categories: hermetic and web-based (Mozgovoy, Kakkonen, & Cosma, 2010). While web-based sources are those that have already been published on the web, hermetic sources are those included in the system’s own database and usually supplied by the clients. The wider the client network of a plagiarism detection tool, the greater its hermetic power. Plagiarism detection tools are not only powerful due to a wide range of texts they can use for comparison, but also because they are capable of using a range of computational detection strategies, including string matching, parsing, the use of a synonym thesaurus, and latent semantic analysis (LSA). While the use of a synonym thesaurus can help identify plagiarism in cases where some of the original wording has been replaced by synonyms, LSA is a computational technique used to identify textual similarities in meaning. However, advanced techniques of natural language processing, such as parsing or LSA, seem to be rarely utilised by plagiarism detection tools, which sets limits to their effectiveness (Mozgovoy, Kakkonen, & Cosma, 2010).

Writers who are determined to plagiarise have been known to develop new strategies aimed at countering the surveillance through anti-plagiarism software. Based on the above computational detection strategies, Maurer, Kappe and Zaka (2006) as well as Kakkonen and Mozgovoy (2010) were able to identify five different types of plagiarism. They are: verbatim copying (copy-pasting), hiding instances of plagiarism (e.g., through poor paraphrasing, deliberate misspelling etc.), technical tricks (inserting characters from a foreign alphabet set, inserting characters in white, insertion of text scanned as image), fake references and “tough plagiarism”, i.e., such that is difficult to detect (using ideas, organisation, references, translation, ghost writers, or medium conversion). Consequently, there is a constant competition between those who engage in plagiarism and those who develop plagiarism detection software.

For all of the above reasons, the effectiveness of anti-plagiarism software in respect of the new computational deception strategies, such as tough plagiarism and conventional copying, becomes a deciding factor in the choice of such a tool. Some examples of plagiarism detection software are programs such as Turnitin, SafeAssignment, AntiPlagiarist, Plagiarism-Finder, SeeSources.com, Sherlock, or WCopyFind. Mozgovoy, Kakkonen and Cosma (2010) conducted an evaluation of some of these tools. One of the main problems identified in some of these packages were false positives, i.e. text that was not plagiarised falsely being flagged for plagiarism. Turnitin and SafeAssignment emerged as the top performance tools, mainly because they did not flag any false positives. SafeAssignment was in fact the best at recognising sources from the Web, but its performance in the hermetic mode was not equally effective. Turnitin on the other hand seemed to have a gap in its Internet search mode, however it seemed quite effective at addressing technical tricks and locating sources in its own database, thus demonstrating what is termed considerable mill strength.

Given the above evaluation, it does not surprise that a number of educational organisations opt for one of the two highest ranking plagiarism detection tools. The organisation hosting the present study happened to be using Turnitin. Like SafeAssignment, Turnitin generates a similarity report. This report does not necessarily represent a measure of plagiarism, but the level to which the examined document is similar to any other documents that the plagiarism detection tool has access to (Walker, 2010). A careful scrutiny of the report by the instructor is required in order to eliminate any possible false positives (Walker, 2010). Depending on the settings, such tools are capable of identifying frequently used multi-word collocations or idioms as cross-textual similarities. However, these are chunks of language that contribute to the native-line quality of expression (Nation, 2001). In addition, asymmetrical quotation marks or quotation marks using unconventional character sets can lead to direct quotation attempts being identified as cross-textual similarity. The fact that direct quotations can be excluded from the comparison is therefore only helpful, insofar as the above is not the case.

Given the technological affordances of plagiarism detection software, an appropriate question to raise would be regarding the best possible use of such software. Three approaches can be identified in the literature, resulting in punitive, deterrent and learning oriented uses of plagiarism detection software respectively (Sutherland-Smith, 2010). The first is based on the assumption that plagiarism is mainly an ethical issue, focusing on how to treat individuals found plagiarising (East, 2009). The second one treats plagiarism as a behaviour problem, whereby the knowledge that plagiarism detection will be used conditions the avoidance of plagiarism. The third approach, which perceives plagiarism detection software as a learning tool, is based on the premise that plagiarism does not have only one cause (Erkaya, 2009; Sharma, 2011; Volkov, Volkov, & Tedford, 2011). It therefore focuses on the learning process, which combines informing the students about the mechanisms of citation, honing the skills of paraphrasing and allowing for experiential learning to take place, with the aid of a plagiarism detection tool. This paper focuses on the latter approach to plagiarism.

Elandera et al. (2010) found that “students with poorly developed authorial identity may be at risk of unintentional plagiarism” (p. 157). Those students were able to develop more confidence in their own writing after intervention which entailed paraphrasing and citing instruction (Elandera et al., 2010). On the other hand, Walker (2010) showed that awareness of the use of anti-plagiarism software by the institution can be a powerful deterrent. Erkaya (2009) however pleaded for a combination of the two approaches. Sharma (2011) suggested a repeated comparison of one’s own manuscript with the wording of the sources. It would seem that such comparison would be best left to the software. For example, Turnitin software can

be programmed to operate in a draft mode. This means that the students' documents will be checked for plagiarism, but not uploaded into the system's database, thus allowing multiple drafts of the paper to be checked, without being cross-referenced with each other.

This software also provides a means of measuring plagiarism rates and their respective decrease. Walker (2010) used this feature to assess the effectiveness of Turnitin as a deterrent. On the other hand, Elandera et al. (2010) and Volkov, Volkov and Tedford (2011) measured the success of their instructional interventions by using soft data (Johns, 1997), namely questionnaires, in a situation in which hard data (Johns, 1997), such as the actual similarity levels would be more helpful. The present study combines the advantages of instructional intervention with both formative feedback and a measure of plagiarism afforded by the Turnitin software, expressed in terms of the percentage of cross-textual similarity.

STUDY AIMS

The main aim of this case study is to investigate the effectiveness of the combined writing instruction and the formative use of Turnitin plagiarism detection software, in comparison with its uses as both a deterrent and a punitive instrument, within the context of English-medium higher education in the Middle East. Since the institution at which the study was carried out is in a country where English is not spoken as the first language, but it is often used in educational contexts, the context and the participants can be described as English as a second language or ESL. The research question reads: Is Turnitin more effective with ESL tertiary students when used as a learning tool than it is in its punitive and deterrent roles respectively? The sections on context and methodology found below explain how the question was addressed.

CONTEXT, TYPE AND ETHICS

The context of enquiry in this study is very much practice oriented and guided by the quest for solutions to a specific educational challenge. It therefore bears some similarity with the context of Walker's (2010) study, in which the researcher examines the effect of plagiarism detection software on the writing of his own students over a period of time. The fact that the investigation described in this paper occurred at one institution places this piece of research within the scope of a case study. Its reflective practice orientation, educational improvement objective, teacher initiative, and naturalistic setting, rather than a strictly controlled environment, position it within the framework of action research (Costello, 2003; McNiff & Whitehead, 2002).

Action research often takes a question from educational practice and raises it to the level of enquiry (McNiff & Whitehead, 2002). While this type of research does not exclude large-scale studies, it is open to small-scale ones, for the benefit of individual teachers who seek to improve their teaching practice (Costello, 2003), such as was the case in the present study. All of the activities described in this paper constituted authentic classroom proceedings which resulted in coursework that was marked accordingly and counted toward the final grade. While the ability to paraphrase and cite sources adequately was included in the marking criteria, the University policy stated that verbatim plagiarism should lead to failure. The timetable of marked assignments is given in Table 1.

The Committee on Human Subjects in Research approved the protocol for this project, which in addition to the application required that all investigators submit a recent human subject protection training certificate. All the activities were carried out with adherence to the ethical standards called for in the Belmont Report, Declaration of Helsinki and Nuremberg Code.

Table 1

The Timetable of Assignments in the Class Type Represented in the Study

Research Proposal	Week 3
Preliminary Outline	Week 5
Preliminary Bibliography	Week 5
Note-Taking (3 sets)	Weeks 7, 9, & 11
Oral Research Report	Week 10
Draft One (development)	Week 12
Draft Two (revision)	Week 13
Introduction & Conclusion	Week 14
Complete Final Paper (Including: Abstract, Title Page, and Reference Page)	Week 15

The complete final paper draft with abstract and reference page was to be 1,800-2,000 words in length, and had to be based on at least 10 source texts, two of which had to be books, four refereed journal articles, and four from any reliable published text-based source, including authored and dated web pages as well as educational or organisation-related websites. In addition to using the library reference tools to locate information sources, the course learning outcomes included specifying a research topic and purpose, as well as planning and conducting secondary research. Summarising, paraphrasing and synthesising from sources, making inferences, and drawing conclusions from several source readings on the same topic were the intended learning outcomes related to the combined reading and writing activities. Finally, the students were expected to be able to evaluate and select source material, develop a viable thesis, organise the paper based on the proposed thesis, as well as to compose and revise their writing in successive drafts.

Samples were provided for each assignment type and the students learned to evaluate and rank sample assignments of different quality. Based on the data provided by students, groups working on related yet independent topics would then plan their respective assignments in class. The development would continue outside the classroom, individually, followed by the in-class peer-review of the draft. The peer-reviewed assignments were further developed at home, individually. Having the draft reviewed by the instructor in one-on-one office hour conferences was encouraged and taken advantage of by a number of students. Revision subsequently took place outside the classroom, after which the assignments were submitted. The submission method was either via email or through Turnitin, depending on the condition as outlined below.

METHODOLOGY

In this study, there are three conditions under which Turnitin was used, each associated with a different section of students enrolled in the same type of writing class, as described above. Turnitin was consequently used as: 1) a punitive measure, 2) a deterrent, and 3) a learning

tool. The first condition involved the use of Turnitin only for the complete final draft of the paper in the final week of the course, without prior warning to the students. The second condition entailed the uploading of three consecutive and subsequently graded paper drafts over three weeks on Turnitin. The third condition involved the uploading of every written assignment on Turnitin, including the proposal, outline, notes and all three drafts of the paper. In addition, under the third condition, Turnitin was used as a learning system to provide formative feedback, from peers as well as the instructor. This on-line feedback was received not as a substitute for, but in addition to the in-class peer-review and one-on-one conferences with the instructor, thus contributing toward a rich learning environment. While both peer-reviews and individual conferences with the instructor were available to the students under the first two conditions as well, Turnitin was not a part of these activities. Under the third condition, the feature which allowed the students to identify the weaknesses in their paraphrasing and quoting strategies and repair them without penalty was the time span between the post-date and the due date. Since the students were allowed to post their assignments on Turnitin several days ahead of the due date, they were able to use the similarity report, which was set up to regenerate once within 24 hours after re-posting, until the due date.

The similarity reports contain two types of information: firstly, the overall percentage of similarity that the text under scrutiny has with the entire repository of texts available to Turnitin, and secondly the percentage of similarity with each text borrowed from. In addition, for the texts in the public domain, access to the source text is provided. Moreover, all cross-textual similarities are highlighted, using colour-coding and numbering according to the source texts. This meant that a student who was able to upload a draft three days in advance of the due date was able to get three consecutive similarity reports, provided that he or she decided to revise the draft once per day, based on the similarity report. More detail is found in the procedure section.

Sample

The convenience sample in this study were 75 second year university students at an English-medium university in the Middle East, enrolled in three English writing classes, originally 25 in each. Each of the three classes was exposed to a different condition. Their first language was Levantine Arabic, while English was one of the additional languages they were conversant with. The age of these multilingual students was between 18 and 20. Forty-one of these were female, while the rest were male. The majors they had declared were engineering (mechanical, electrical, and civil), health sciences, business, architecture and computing. Their English proficiency scores were the equivalent of 550 or higher on paper based TOEFL. However, only 52 of those participants submitted all assignments and participated in all of the classroom activities with 22 of those under condition 1 (punitive measure), 14 under condition 2 (deterrent) and 16 under condition 3 (learning tool). These numbers are reflected in the statistical analyses of data presented in the results.

Procedure

As explained in the objectives section, each group was exposed to one of the three conditions. While all of the participants in this study received comparable instruction in paraphrasing, quoting, summarising and acknowledging sources, only one of the groups integrated Turnitin in their learning activities. These ESL students were encouraged to run their drafts through the Turnitin software prior to submission and then use the Turnitin similarity report to

improve their paraphrasing and quoting skills. Before grading, several revisions were possible, until the students were satisfied that any similarity which could be construed as plagiarism had been removed. In the process, the instructor and peers could also be consulted without penalty.

The procedure described above was compared with the two scenarios examined in previous research (Walker, 2010), namely the use of software as a detector and deterrent. As pointed out above, these two conditions included the use of Turnitin to upload 1) the final assignment draft only, 2) all assignment drafts for grading purposes. Only the third group was invited to upload all notes and drafts after explicit instruction in the use of Turnitin to improve paraphrasing skills, and with numerous opportunities to revise before grading.

The Turnitin assignment settings were carefully selected, in order to exclude citations and collocations of up to 4 words in length from the textual comparison. To eliminate other false positives, students were required to submit their title pages and reference lists separately, as these can be flagged by Turnitin. For the purpose of this study, the final drafts of the last course assignment were used as the main means of comparison between groups. These were examined by two independently working instructors, who found that due to the above measures, accidental cross-textual similarity could be excluded. Most of the reported instances of similarity did not seem to constitute copy-pasting of the entire paragraphs of borrowed text, but could be described as patchwriting (Amsberry, 2010), indicating that poor paraphrasing skills might be the most common cause of textual borrowing. Only three instances of verbatim copying from sources and one instance of purloining or borrowing from classmates (Walker, 2010) could be identified under the punitive condition, and none under the other two conditions. Therefore, it was decided that the overall percentage of cross-textual similarity for each student paper could be adopted as an adequate measure of plagiarism. These percentages were entered into anonymised spreadsheets and used as raw data.

Limitations

The limitations of the study are typical of those encountered in action research in general. One of the limitations of the present study is its convenience sample, which is by definition not randomised, and consequently makes it difficult to generalise from the findings. In addition, the sample is small, further affecting the possibility of generalising from data, which is what this study has in common with case studies in general.

RESULTS

To sum up, each of the three groups wrote a multiple draft assignment, but only the deterrent and learning tool groups posted all of the drafts on Turnitin, while the punitive use group posted only the final draft on Turniti, without advance notice. In addition to all paper drafts, the learning group also posted notes in several drafts on Turnitin. While all groups received similar instruction, the learning group was also the only group which continuously included Turnitin to test hypotheses about adequate paraphrasing of sources. The outcome of the procedure was measured using Turnitin similarity reports and expressed as a percentage of cross-textual similarity. As indicated before, Turnitin reports indicate the level of similarity between the analysed text and any other texts accessible to Turnitin. Consequently, the three sets of results represent the similarity measures for the three conditions: 1) only the final draft of the assignment submitted to Turnitin without prior notice (punitive use), 2) each of the

three drafts was submitted to Turnitin (deterrent use), 3) not only each draft was submitted to Turnitin, but every stage in assignment drafting, with the students being given a grace period to check their similarity levels for each stage and revise, if necessary, several times (learning tool use). In addition, Turnitin was used to enrich the learning environment by assisting with the choice of an original research question on the one hand, and serving as an avenue for the delivery of feedback on the other. The results of the final drafts of the last assignment for each group are analysed in the following section.

Data Analysis

An initial one-way analysis of variance (ANOVA) for independent samples, performed using the VassarStats online software package (Lowry, 2012), indicated that the participants differed in their performance on the three test conditions (data sets 1, 2 & 3). The values for these data sets are summarised in Table 2, whereas Table 3 contains an ANOVA analysis summary, with the key indicator of difference. This indicator, $F = 7.24$ ($p < .0001759$), suggests that the overall difference among the groups was statistically significant.

Table 2
Data Summary for the Three Conditions

<i>Data Summary</i>						
	Samples					
	1	2	3	4	5	Total
N	22	14	16			52
$\sum X$	508	228	17			753
Mean	23.0909	16.2857	1.0625			14.4808
$\sum X^2$	24474	6366	41			30881
Variance	606.8485	204.0659	1.5292			391.7055
Std.Dev.	24.6343	14.2852	1.2366			19.7916
Std.Err.	5.252	3.8179	0.3091			2.7446

Table 3
ANOVA Analysis Summary for the Three Conditions

standard weighted-means analysis					
ANOVA Summary Independent Samples k=3					
Source	SS	df	MS	F	P
Treatment [between groups]	4557.3679	2	2278.684	7.24	0.001759
Error	15419.6128	49	314.686		
Ss/Bl					
Total	19976.9808	51			

Table 4
Tukey's HSD Test for the Three Conditions

Tukey HSD Test

```
HSD[.05]=14.82; HSD[.01]=18.73
M1 vs M2  nonsignificant
M1 vs M3  P<.01
M2 vs M3  P<.05
```

Tukey's HSD (honestly significant difference) test for multiple comparisons between groups of results was subsequently used (Table 4). While there was no significant difference between conditions 1 and 2, i.e. the punitive and deterrent uses of Turnitin respectively, condition 1, i.e. the punitive use yielded significantly higher similarity levels than condition 3, i.e., the learning tool use ($q = 18.73$, $p < 0.01$). Similarly, condition 3, i.e. the learning tool use indicated significantly lower similarity levels than condition 2, i.e. the deterrent use ($q = 14.82$, $p < 0.05$).

In order to have some means of comparison with Walker's (2010) study, which measured the success of the use of Turnitin as a deterrent, a paired t-test was performed using the first two datasets from the second condition. It indicated that the participants differed significantly ($t = 6.9913$, $p < 0.0001$) in their performance on the first and second assignment uploaded to Turnitin. Table 5 below contains the mean similarity percentage values and the standard deviation for each of the three drafts under condition 2. The values indicate the progressive decrease in cross-textual similarity with the increased awareness of anti-plagiarism software.

Table 5

Mean and Standard Deviation for the Three Drafts under Condition 2

Condition 2	Paper draft 1	Paper draft 2	Paper draft 3
Mean	23.52	16.29	9
Standard deviation	19.25	14.29	8.5

For condition 3 however, there were 10 consecutive assignments, all of which were posted on Turnitin. For better comparison with conditions 1 and 2, Table 6 below includes only the initial and the final assignment values, expressed again in terms of the mean and standard deviation.

Table 6

Mean and Standard Deviation for the Initial and Final Assignment under Condition 3

Condition 2	Proposal (Initial)	Paper draft 3 (Final)
Mean	8.95	1.06
Standard deviation	7.36	1.23

However, when working with small groups, such as is the case in this study, the median may provide a more appropriate measure. For this reason, the Table 7 below includes a comparison of the three conditions using the median, the value in the middle of the measurement array. For condition one, only the value for final paper draft was available, while conditions 1 and 2 are compared using the three paper drafts as a criterion.

Table 7

Similarity Percentage Median for Each of the Three Paper Drafts per Condition

Condition	Paper draft 1	Paper draft 2	Paper draft 3
1			17%
2	5%	2%	3%
3	1%	1%	0%

DISCUSSION

The data sets show a tendency toward decrease in similarity among the three different uses of Turnitin. However, this decrease is not significant between the first group, which had to upload only the final draft on Turnitin, without prior warning, and the second group, which knew from the beginning of the course that all three graded drafts were to be uploaded on Turnitin. While this might seem to be a departure from Walker's (2010) findings, according to which most plagiarism types dropped significantly from the first assignment to the next, analysis on another dataset is consistent with Walker's findings. Consequently, in the second group, which uploaded all of the graded drafts of each assignment on Turnitin, there is a significant drop in the similarity rates on the second assignment, followed by a less pronounced decrease on the third assignment. Thus, the use of Turnitin as a deterrent would appear to be as effective as in Walker's study. However, it does not seem to be significantly more effective than the use of Turnitin as a punitive measure in the present study.

Despite some similarities, there are considerable differences between the present study and the one conducted by Walker (2010). First and foremost, Walker's study only used Turnitin as a deterrent, and not as a learning tool. Second, his study was conducted on a much larger sample, precisely eight times larger, taking only repeated measures into account, while the present study was mainly based on independent samples. The latter only deviated from this principle in the part where the comparison with Walker's study was made, i.e. measuring the decrease in cross-textual similarity. Finally, in addition to breaking down the results by sub-category of the sample (e.g., gender, nationality, mode of study), Walker's study also clustered the instances of plagiarism around several types, depending on the degree and source of textual appropriation. This is not the case in the present study.

The limitations of the present study are however akin to those of Walker's (2010) study, in the sense that the sample is an unrandomised convenience sample, which makes generalisations from the findings less likely. However, a major difference is that the sample is smaller than Walker's sample, further restricting the generalizability of the conclusions. Like Walker's study, the present study is limited to a particular educational context, and can only be used indicatively. There is nonetheless one significant difference between the constructs measured in the two studies. While Walker admits that by developing a typology of plagiarism instances the onus was on the lecturer to categorise each instance, thus adding a subjective dimension to the study, the present study only takes into account such Turnitin output which reports unclassified similarity, identified by two independent instructors as patchwriting, which reduces the subjectivity of the study to some extent.

Given the results, it is the third group, which uploaded every stage in the writing process to Turnitin, that shows minimum levels of cross-textual similarity. While condition 1 (only final draft of the last assignment subjected to Turnitin analysis) appeared to be the least effective, condition 2 (final drafts of three consecutive assignments uploaded) yielded lower levels of cross-textual similarity. Within sub-sample however, the use of Turnitin as a deterrent under this condition seemed successful, as there was a statistically significant difference between the first and the second assignment. However, the between sub-samples decrease of cross-textual similarity rates between conditions 1 and 2 was not statistically significant, suggesting that the sheer awareness of anti-plagiarism software was in this case not necessarily more effective than the punitive use of the software. On the other hand, there was a statistically significant difference between the third condition, under which Turnitin was used as a learning tool, and both other conditions.

Another interesting detail that sets apart the third group of results (learning tool) is the fact that 0 was the mode value of the data set, accounting for half of the similarity percentages in this group. What can be assumed with a high degree of likelihood is that 0% similarity means the absence of plagiarism, at least from such sources that are available to Turnitin. Hard copy text books and similar sources are only detectable if they have been either cited or copied from before, within the Turnitin database.

This raises the question whether the students have truly learned how to avoid plagiarism or just how to avoid being caught. However, the gradual development of thoughts and paragraphs, over 10 stages, from research proposal over outline, notes and multiple drafts to the final paper, a great deal of which happened in class or in one-on-one conferences with the instructor, provided a reasonable guarantee that that the texts uploaded were genuinely authored by the student participants. As proposed by Erkaya (2009), both the paper development and sometimes the posting on Turnitin happened under the watchful eye of the

course instructor. The end-product of the writing process was a portfolio with the printouts of all the readings, notes on the margins of the readings, underlined and highlighted text of the original, in addition to both the electronically submitted assignments and notebooks or any scratch paper evidence of process writing, serving as evidence of thought processes and gradual text development.

The grace period between the post-date and the due date, during which each student was able to obtain one similarity report per day by resubmitting the assignment, was taken advantage of in the total of 96% of all the opportunities given to each individual student in this group, which suggests that the students were really interested in learning how to improve their paraphrasing skills, as poor paraphrasing was the most common cause of cross-textual similarity in this group. Consequently, it would appear that using anti-plagiarism software as a learning tool, rather than just as a detector or deterrent is an effective way of helping ESL student writers find their authentic voice.

Implications for Practice

From the perspective of action research, this project had important implications for the teaching practice within the same institution. The teacher researcher who conducted this study felt the need to share the outcome with colleagues. With the support of the administration, this instructor prepared workshops on the use of Turnitin as a learning tool for other writing faculty across campuses. The workshops were well attended and received, while the colleagues are expected to report their results in the unit meetings.

CONCLUSIONS AND RECOMMENDATIONS

This action research study examined three different ways in which Turnitin plagiarism detection software can be used in English-medium tertiary writing classes within the Middle East context, as: 1) a punitive measure, 2) deterrent and 3) learning tool. The findings suggest that in this particular case study, the third condition, in which Turnitin was used as a learning tool, to test the students' hypotheses about effective paraphrasing and make adequate adjustments, was the most effective way. The cross-textual similarity percentage, used by Turnitin as plagiarism indicator, dropped to zero in one half of the final draft reports, without a compromise in text accuracy and readability, indicating that this particular group of students was able to learn how to paraphrase more effectively, while avoiding plagiarism, with the formative use of plagiarism detection software as one of their learning resources.

Apart from the use of Turnitin in independent study mode, the context of this outcome included both classroom interactions with the instructor as well as peers and individual consultations with the instructor. Consequently, it was not expected of students to learn from Turnitin completely autonomously. It is suggested therefore that, if this practice is used, it be imbedded in a writing class which covers detailed instruction in paraphrasing. The independent Turnitin aspect of such a course can only be useful if the students already have a solid grasp on effective paraphrasing methods, and have a community, i.e. classmates and instructor, to get feedback from on ideas for improvement. While Turnitin can clearly demonstrate the shortcomings of their hypotheses about effective paraphrasing, it cannot help with the choice of words or grammatical structures. In addition, the instructor really needs to make sure that with all the changes of wording the meaning of the original has remained preserved. As proposed by Erkaya (2009), Turnitin should be used repeatedly, in the

formative mode, under the supervision of the instructor. It is under those conditions that its use was successful in this study.

ACKNOWLEDGMENTS

The author is indebted to the reviewers, whose comments have been exceedingly helpful.

REFERENCES

- Amsberry, D. (2010). Deconstructing plagiarism: International students and textual borrowing practices. *The Reference Librarian*, 51, 31-44.
- Anderson, J., & Poole, M. (1994). *Thesis and assignment writing*. Brisbane: John Wiley & Sons.
- American University of Sharjah (2007). *Plagiarism statement*. Course materials: Sharjah.
- Ballard, B., & Clanchy, J. (1984). *Study abroad: A manual for Asian students*. Kuala Lumpur: Longman.
- Basharina, O. K. (2007). An activity theory perspective on student-reported contradictions in international telecollaboration. *Language Learning & Technology*, 11(2), 82-103.
- Bloch, J. (2007). Abdullah's Blogging: A Generation 1.5 Student Enters the Blogosphere. *Language Learning & Technology*, 11(2), 128-141.
- Costello, P. J. M. (2003). *Action research*. London: Continuum.
- East, J. (2009). Judging plagiarism: A problem of morality and convention. *Higher Education*, 59, 69-83.
- Elandera, J., Pittamb, G., Lusherc, J., Foxd, P., & Paynee, P. (2010). Evaluation of an intervention to help students avoid unintentional plagiarism by improving their authorial identity. *Assessment & Evaluation in Higher Education*, 35(2), 157-171.
- Erkaya, O. R. (2009). Plagiarism by Turkish students: Causes and solutions. *Asian EFL Journal*, 11(2), 86-103.
- Evans, F. B., & Youmans, M. (2000). ESL writers discuss plagiarism: The social construction of ideologies. *Boston University Journal of Education*, 182(3), 49-65.
- Hyland, F. (2001). Dealing with plagiarism when giving feedback. *ELT Journal*, 55(4), 375-381.
- Johns, A. M. (1997). *Text, role and context. Developing academic literacies*. New York: Cambridge University Press.
- Kakkonen, T., & Mozgovoy, M. (2010). Hermetic and web plagiarism detection system for student essays – An evaluation of the state of the art. *Journal of Educational Computing Research*, 42(2), 135-159.
- Lankamp, R. (2009). ESL student plagiarism: Ignorance of the rules or authorial identity problem? *Journal of Education and Human Development*, 3(1), 1-8.
- Lowry, R. (2012). VassarStats: Website for statistical computation. Retrieved from <http://faculty.vassar.edu/lowry/VassarStats.html>
- Lund, J. R. (2004). Plagiarism: A Cultural Perspective. *Journal of Religious & Theological Information*, 6 (3/4), 93 – 101.
- Maurer, H., Kappe, F., & Zaka B. (2006). Plagiarism – A survey. *Journal of Universal Computer Science*, 12(8), 1050-1083.
- McNiff, J., & Whitehead, J. (2002). *Action research: Principles and practice*. London: Routledge.

- Mozgovoy, M., Kakkonen, T., & Cosma, G. (2010). Automatic student plagiarism detection : future perspectives. *Journal of Educational Computing Research*, 43(4), 511-531.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Reingold, R. & Bartz, L. (2011). An institutional code of ethics – A response to attitude of Israeli teachers' education college students towards academic plagiarism, *US-China Education Review*, 8(5), 589-598.
- Sharma, R. (2011). A step-by-step guide to students: How to avoid plagiarism. *Journal of Education Research*, 4(2), 143-153.
- Shi, L. (2006). Cultural backgrounds and textual appropriation. *Language Awareness*, 15(4), 264-282.
- Sutherland-Smith, W. (2010). Retribution, deterrence and reform: The dilemmas of plagiarism management in universities. *Journal of Higher Education Policy and Management*, 32(1), 5-16.
- Volkov, A., Volkov, M., & Tedford, P. (2011). Plagiarism: Proactive prevention instead of reactive punishment. *e-Journal of Business Education & Scholarship of Teaching*, 5(2), 22-35.
- Walker, J. (2010). Measuring plagiarism: Researching what students do, not what they say they do. *Studies in Higher Education*, 35(1), 41-59.
- XJTLU (2012). Code of practice on assessment Retrieved from <http://www.xjtlu.edu.cn/en/component/content/article/147-learning-a-teaching/policies-and-regulations/653-code-of-practice-on-assessment.html>

ⁱ The author of this paper has no affiliation with Turnitin.