EFL teachers’ use of plagiarism detection software: A usability study

Abdul Syahid
abdal.syahid@iain-palangkaraya.ac.id
Department of English Language Education, Faculty of Teacher Training and Education, Institut Agama Islam Negeri Palangka Raya, Indonesia
Jalan G. Obos Kompleks Islamic Center, Palangka-Raya 73112, Indonesia

Abstract

EFL teachers sometimes struggle harder to prevent their students’ committing plagiarism because the students’ limited EFL proficiency could be a cause of plagiarism. To prevent plagiarism, one of the methods widely employed is the use of plagiarism detection software such as Plagiarism Checker X. To assess how useful, easy to learn, easy to use, and satisfying the software was for them, 30 EFL teachers took anonymously an online 30-item questionnaire of 7-point Likert scale after they had downloaded and run a free version of the software. High mean agreement ratings across the four dimensions were revealed as 28 items were highly positively perceived between 6 (agree) and 7 (strongly agree) and two items were rated between 5 (slightly agree) and 6 (agree). The study showed that the software was useful for checking similarities between texts, easy to learn and use and could satisfy the teachers. They could arm themselves with the software to prevent plagiarism. As one of the first attempts to evaluate the usability of text-matching software from EFL teachers’ perspective, this study has research and educational implications as it highlights avenues for preventing plagiarism via the software usability evaluation.

Keywords: usability study; plagiarism detection software; plagiarism prevention; teachers of English as a foreign language

INTRODUCTION

In the academic arena, plagiarism has long been public enemy number one. Voices concerning the prevalence of one of the academic misconducts in which, intended or not, someone or a group of people used any other persons’ scientific or artistic work without giving proper credits in both published and unpublished academic or artistic work (Minister of National Education, 2010) have been raised across times as noted by Sadeghi (2019). This type of academic misconduct has been committed by not only students from elementary school (Ding et al., 2014) to graduate ones (Ogawa & Sundberg, 2019) but also professors (Siaputra & Santosa, 2016) in developing (Rohwer, Wager, Young, & Garner, 2018) and developed countries (Sarah Elaine Eaton, 2017). The “disease” has done some irreparable harm to the plagiarists themselves, other people, and community (‘Teddi’ Fishman, 2016, p. 14) such as the loss of an academic
career, institutional status, and people’s lives (Turnitin, 2012). Because of the damages it can produce, plagiarism must be prevented from the very beginning such as in elementary schools as called by Wan and Scott (2016).

The early intervention is a must since plagiarism has a wide spectrum of causes and types, among others. In their overview of factors leading to plagiarism, Eaton et al. (2017) listed individual factors such as laziness and contextual ones such as limited attention given to plagiarism cases at an institutional level. Some plagiarism cases resulted from the lack of writing skills and knowledge about research rather than plagiarizing on purpose to get higher scores, especially in some places where plagiarism (more precisely, memorizing and recalling facts) was culturally acceptable (Glendinning, 2016). Regrettably, the dispute over the definitive definition of plagiarism (Carroll, 2016) and how to deal with the plagiarism cases (Babaii & Nejadghanbar, 2017) for decades have been unsuccessfully settled. It could explain why students have limited understanding of the very concept of plagiarism as one of the plagiarism causes.

In terms of forms, some of the plagiarism types by intent, as Turnitin (2016) reported, cannot be easily recognized and need more careful identification such as “404 Error” (p. 16) by which an author cites sources but write no proper information. When some plagiarism scenarios were presented, even 15% of 681 lecturers from Higher Education Institutions in over 20 European countries were not sure whether or not those were plagiarisms and 4% of them found the scenarios were not plagiarism (Glendinning, 2016). Unfortunately, the what and how of plagiarism are clearly not beyond dispute yet. Viewed from the complexities, it could be understood why plagiarism prevention or proactive approach is now given priority over punishment or reactive one.

The plagiarism prevention is every effort aimed at reducing and stopping plagiarism contained in a manuscript before the final form is submitted to a teacher/examiner, presented in a seminar, or published in a certain academic journal. In other words, it “prevents plagiarism from happening rather than detecting it after the event” (Radcliffe & Rudolph, 2008, p. 13). In the context of Indonesian Higher Education Institutions, for example, the plagiarism prevention ranges from the implementation of institutional code of conduct and citation style to an originality statement for each academic paper (Minister of National Education, 2010). After over 5 years, a more operational preventive intervention was taken by the Ministry of Research, Technology, and Higher Education through General Director of Science Technology and Higher Education Resources (Mukti, 2016) underlined the use of plagiarism detection software for the promotion for Associate Professor and (full) Professor. The following year the Ministry of Religious Affairs through General Director of Islamic Education (2017, p. 6) also highlighted the use of software for written academic work by students and teaching staffs in State Islamic Higher Education Institutions and put a limit on 25% similarity tolerance for undergraduate students and 20% for graduate ones. Similarly, a lot of Higher Education Institutions not only in Britain (Brown & Janssen, 2017), Germany (Sattler, Wiegel, & Veen, 2017), and America (Levine & Pazdernik, 2018) but
also around the world (Meo & Talha, 2019) have used such software as one of the plagiarism detection and prevention methods.

The software, either desktop-based such as Plagiarism Checker X (Team Plagiarism Checker X, 2019) or web-based such as PlagScan (The Team of PlagScan, 2019), is used to compare a manuscript submitted to the system against other papers in a certain database. The database consisting of “websites, millions of periodicals, books, and archives of assignments and documents” (Sun, 2013, p. 265). Based on the comparison, the (more appropriately named) text-matching or similarity detection software releases reports consisting of the similarity percentage, whose range is 0% to 100%, of matching or similar texts in the database (“The Similarity Report explained,” n.d.).

In its relevance to the software adoption in academic institutions, much of work on the effectiveness of text-matching software has been carried out. Shang (2019) found that the students who were aware of their institution’s use of a text-matching tool had lower plagiarism percentage in their papers than those who were not. In addition, seeing their similarity reports and understanding the comparison of their papers with other papers in the web as highlighted in the reports, the students had lower plagiarism rate after the use of similarity checking software as a formative writing tool (Halgamuge, 2017). Interestingly, some students were afraid of submitting their papers to the software (p. 9). After integrating the similarity checking software with some plagiarism reduction programs such as discussing socio-moral aspects of plagiarism and learning paraphrasing techniques, Kashian, Cruz, Jang, and Silk (2015, p. 252) consider the software use by the students “a self-monitoring mechanism.” The color-coded similarities in the report would drive the students to review their work and rewrite it using more accurate citation and paraphrase techniques. The results have further strengthened the confidence in the use of software in the plagiarism prevention.

The more plagiarism increases in the era of internet, the more text-matching software is extensively used. Countless text matching software systems has been marketed as plagiarism detection one. Nearly 30 applications were thoroughly tested by Weber-Wulff (2016). More recently, Kamzola & Anohina-Naumeca (2020) compared 16 systems available on the market including Plagiarism Checker X (Team Plagiarism Checker X, 2019). The software offers some features to teachers in their day-to-day activities, especially those in the field of L2 teaching. For instance, it can also be used for texts written in non-Roman alphabets such as Arabic ones (Hussein, 2015) and to check similarity between a paper written in Macedonian and its translated version of English with the highest similarity rate of the three similarity detection tools in Spiroski (2016). In addition, it was used by one of the science journals indexed in Scopus but the most cost effective of three programs in Jain, Das, and Garg (2016). If the software is found helpful, the price affordable price could not hinder its personal and institutional adopts. Interestingly, as a desktop based tool, even in the free version, it can be used offline to check not only two but multiple documents so it can be can be seen how two or more documents have certain similarities. The availability of its free version, features, cost,
and the facts that it has been used by some internationally reputable journals could justify the choice of software in this study.

However, as similarity does not necessarily mean plagiarism, in the context of day-to-day teaching arena, it is the teacher who decides “whether or not a text parallel indeed constitutes plagiarism” (Weber-Wulf, 2016, p. 635). Using the software, teachers who always serve in the front-line of the wars against plagiarism could arm themselves to prevent one of the ethical breaches.

As previously noted, the potential of such software for preventing plagiarism in a school setting has been examined in a growing body of literature. However, not much work could be found on how the teachers use the software, more specifically on how the software is viewed from the teachers’ perspectives. In their study on the usability of a learning and plagiarism avoidance tutorial system integrating two similarity checking tools, i.e. Glatt Plagiarism Program and Turnitin, Liu, Lo, and Wang (2013) offer no explanation for the software usability in particular but the tutorial system in general. Even after taking part in 7 studies on the effectiveness of more than 20 text-matching systems in reporting plagiarism, Weber-Wulf (2016) still overlooks the software usability in details. Another study could better be conducted to enhance our understanding of how well the teachers learn and run the software to detect similarity along with use its generated report as well as how satisfied they are with the application.

Moreover, fewer researchers have addressed how the software was perceived by teachers of second/ foreign language (L2) in terms of usefulness, ease of use, ease of learning, and satisfaction. The L2 teachers might face greater challenges because the level of L2 proficiency could result in plagiarism. Whereas those who have low L2 proficiency, with their limited understanding on the source texts, would not know what to include in their own work while, those who have mid L2 proficiency, even with better understanding on the source texts, have to struggle to paraphrase (Wolfersberger, 2018). In another study, finding that the lower the English as a foreign language (EFL) proficiency level Vietnamese students had, the higher the plagiarism level they committed, Perkins, Gezgin, and Roe (2018) conclude that plagiarism incidences by EFL learners could be reduced through an intervention aimed at improving their English academic skills. The use of text-matching software as pedagogical approach by L2 teachers is, however, poorly understood.

Driven by the lack of a lack of information about the use of similarity checking software by L2 teachers, this study examined how the usability of software was perceived by L2 teachers, EFL ones in particular. In his seminal work, Lund (2001) found that users evaluated software, hardware, services, and user support materials using four dimensions, i.e. Usefulness, Ease of Use, Ease of Learning, and Satisfaction. Even after almost two decades, the questionnaire he developed is still one of the latest and most frequently used instruments to measure subjective reactions to the product usability of a product, e.g. in Graham et al. (2019). This study therefore investigated how the Usefulness, Ease of Use, Ease of Learning, and Satisfaction dimensions of Basic version of Plagiarism Checker X were perceived by EFL teachers and which
dimensions could motivate them to adopt the software in the EFL teachers’ day-to-day combat against plagiarism.

**METHOD**

The participants, instrument, and procedure in this study were the same as in Syahid (2018) but with a different tool to review. The 30 participants took part in this study when joining the second phase of professional development program for EFL teachers held by a language center of the only private university in East Kotawaringin district of Central Kalimantan province, Indonesia. The duration between the first and second phases of the programs was one month.

In terms of the instrument, reusing the same instrument was on the basis of the considerations previously taken into account. That the participants had used it in a previous study could help the them feel more comfortable with completing it. Many researchers such as Dantas et al. (2017) and Gao, Kortum, and Oswald (2018) using the questionnaire reported high reliabilities, i.e. .80 and .98 respectively. In this study, the usefulness, ease of use, ease of learning, and satisfaction subscales of the USE questionnaire as a whole showed an excellent reliability, Cronbach’s 𝛼 = .98. Whereas the 8 items of Usefulness, 11 items of Ease of Use, and 4 items of Ease of Learning also had excellent reliabilities, Cronbach’s 𝛼 > .95, the 4 items of Ease of Learning had a good reliability, Cronbach’s 𝛼 = .85.

Before completing anonymously the Lund’s (2001) seminal questionnaire of Usefulness, Satisfaction, and Ease of use (USE) hosted in a Moodle-based teacher training, i.e. https://syahid.gnomio.com/ in a maximum of 10 minutes, the participants had been guided to download, installed, and run the free version of Plagiarism Checker X (2019). After downloading and install the software, they were asked to run three key functions of the program. They were Online Plagiarism, Side-by-Side Comparison, and Bulk Search (with three comparison methods: one to many, many to many, and cross comparison). The participants were also told how to exclude some web addresses from search.

By submitting the data to IBM SPSS Statistics Version 22 (IBM Corp, 2013) and using descriptive statistics to describe the data, it could be found how the four dimensions were separately and collectively perceived by the participants. The analysis was loosely based on Field (2017).

**FINDINGS AND DISCUSSION**

To address the usability of the Basic version of Plagiarism Checker X as a tool for checking similarity, the responses to the first and second questions would be discussed. In response to the question about how the usefulness, ease of use, ease of learning, and satisfaction dimensions of Basic version of Plagiarism Checker X were perceived by the participants, Figure 1 shows that all of the 30 participants showed a relatively high level of agreement, i.e. between 6 (agree) to 7 (strongly agree) to all of the subscales.
Regarding the descending order of agreement, it can be seen that the Ease of Learning subscales had the highest mean agreement rating, $M = 6.57, SD = 0.50$.

Importantly, all of the participants found the software not only useful for finding/detecting similarities between texts submitted to the systems but also easy to use as indicated by the same agreement rating, $M = 6.47, SD = 0.51$. Even though the satisfaction subscale had the lowest mean agreement one, all of the participants still considered the tool very satisfying and well recommended, $M = 6.17, SD = 0.59$.

![Figure 1. Usability of Basic version of Plagiarism Checker X](image)

The response to question on which usability dimensions that could motivate them to adopt the software in the EFL teachers’ day-to-day combat against plagiarism were all of Ease of learning, Usefulness, Ease of Use, and Satisfaction dimensions of which the participants were satisfied. Regarding the software’s ease of learning, Table 1 shows each question with its average score. It is apparent that three out of 4 items of Ease of Learning subscale had the highest mean agreement rating, $M = 6.57, SD = 0.5$. Even for the lowest mean score in this subscale, the software could score very well with an average of 6.37. The participants thus highly agreed with the software’s ease of learning subscales.

That the Ease of Learning subscale was rated the highest strongly suggests that the participants who have never used the software could run the software skillfully once it was learnt (Lund, 2001). In the context of education, the ease of learning dimension such as learnability and memorability could encourage the EFL teachers to use the software as mentioned by Faria, Pavanelli, and Bernarde (2016) when using the same instrument to assess the usability of a web-based educational system.
Table 1. Ease of Learning items

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>I learned to use it quickly.</td>
<td>6.57</td>
</tr>
<tr>
<td>I easily remember how to use it.</td>
<td>6.57</td>
</tr>
<tr>
<td>It is easy to learn to use it.</td>
<td>6.57</td>
</tr>
<tr>
<td>I quickly became skillful with it.</td>
<td>6.37</td>
</tr>
</tbody>
</table>

The next subjective factor that could influence the teachers’ adoption of the software is its perceived Usefulness. The participants showed a high level of agreement ratings for all of them items. As can be seen in Table 2, the mean agreement ratings were in the range of 6.17 to 6.47. In contrast with what Lund (2001) found, even though the participants was introduced to the software in such a limited time, they were not more variable in their Usefulness ratings. This can be attributed to the unlimited offline functions offered by the free version of the text matching tool.

Table 2. Usefulness items

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>It helps me be more effective.</td>
<td>6.47</td>
</tr>
<tr>
<td>It helps me be more productive.</td>
<td>6.47</td>
</tr>
<tr>
<td>It is useful.</td>
<td>6.47</td>
</tr>
<tr>
<td>It gives me more control over the activities in my life.</td>
<td>6.47</td>
</tr>
<tr>
<td>It makes the things I want to accomplish easier to get done.</td>
<td>6.47</td>
</tr>
<tr>
<td>It saves me time when I use it.</td>
<td>6.27</td>
</tr>
<tr>
<td>It meets my needs.</td>
<td>6.47</td>
</tr>
<tr>
<td>It does everything I would expect it to do.</td>
<td>6.17</td>
</tr>
</tbody>
</table>

This study reveals that in its Online Plagiarism checking function the free version of Plagiarism Checker X limited documents/ assignments, paragraphs, or webpages) to 150 words maximum for a comparison with nearly 20 billion published pages in the Google or Bing database (Team Plagiarism Checker X, 2019). Fortunately, in its Offline Plagiarism checking one, the desktop application did not place a limit. In Side-by-Side comparison, the participants could detect similarities between two documents in two windows and find same contents effortlessly. In addition, the feature of Bulk-Cross Comparison allowed the participants to determine originality in one document to many documents, a group of documents to another group, and multiple documents against each other within the local database.

Side-by-Side Comparison and Bulk Search were very useful for detecting similarities between documents submitted by students. Even though the maximum number of documents has not thoroughly examined in this study, one of the participants run the feature of Bulk-Cross Comparison to determine originality in 32 documents against each other within a folder in the participants’ notebook and could detect similarities between three students’ papers. It is not surprising that usefulness dimension of the Basic version was well rated in this study.
In this usability study, the subjective factor having the same mean agreement rating as Usefulness subscale was the Ease of Use one. The same level of agreement between Usefulness and Ease of Use in this study is barely distinguishable from Lund (2001) who found that the improvement in one rating improved another one. Table 3 shows that all of the mean agreement ratings but “the system consistency” item were between 6 (agree) to 7 (strongly agree).

Table 3. Ease of Use items

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is easy to use.</td>
<td>6.47</td>
</tr>
<tr>
<td>It is simple to use.</td>
<td>6.47</td>
</tr>
<tr>
<td>It is user friendly.</td>
<td>6.47</td>
</tr>
<tr>
<td>It requires the fewest steps possible to accomplish what I want to do with it.</td>
<td>6.17</td>
</tr>
<tr>
<td>It is flexible.</td>
<td>6.47</td>
</tr>
<tr>
<td>Using it is effortless.</td>
<td>6.47</td>
</tr>
<tr>
<td>I can use it without written instructions.</td>
<td>6.47</td>
</tr>
<tr>
<td>I don't notice any inconsistencies as I use it.</td>
<td>5.97</td>
</tr>
<tr>
<td>Both occasional and regular users would like it.</td>
<td>6.47</td>
</tr>
<tr>
<td>I can recover from mistakes quickly and easily.</td>
<td>6.47</td>
</tr>
<tr>
<td>I can use it successfully every time.</td>
<td>6.17</td>
</tr>
</tbody>
</table>

The least mean agreement rating in this study was received by the Satisfaction subscale. However, as can be seen in Table 4, the rating remains high with the range of 5.2 to 6.37. Only “I feel I need to have it” M = 5.20, SD = .97, had the mean agreement ratings between 5 (slightly agree) and 6 (agree). Overall, the participants enjoyed using the software.

Table 4. Satisfaction items

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am satisfied with it.</td>
<td>6.17</td>
</tr>
<tr>
<td>I would recommend it to a friend.</td>
<td>6.37</td>
</tr>
<tr>
<td>It is fun to use.</td>
<td>6.37</td>
</tr>
<tr>
<td>It works the way I want it to work.</td>
<td>6.37</td>
</tr>
<tr>
<td>It is wonderful.</td>
<td>6.37</td>
</tr>
<tr>
<td>I feel I need to have it.</td>
<td>5.20</td>
</tr>
<tr>
<td>It is pleasant to use.</td>
<td>6.17</td>
</tr>
</tbody>
</table>

The study was limited to participants at one location and only one tool was assessed. Further studies are needed to determine whether these findings could be applied to the Pro and Business versions of Plagiarism Checker X selling for around US$40 (one Windows based computer) and US$148 (up to 5 computers) equipped with some extended functions as detailed by Team Plagiarism Checker X (2019). The participants included only some EFL teachers so result may differ for other participants with different backgrounds or for the users of software in general. The difference results
could result from the participants having longer exposure to the software. Nevertheless, this study could provide valuable information about the perceptions of the usability of a text-matching tool from the viewpoint of EFL teachers that adds to the body of knowledge about the use of such as one of plagiarism prevention methods.

CONCLUSION

The set of questions about Usefulness, Ease of Use, Ease of Learning, and Satisfaction aims to measure the experience of the EFL teachers with one of text-matching tools, i.e. the Basic version of Plagiarism Checker X. All usability dimensions of the software were highly positively rated in this study. None of the dimensions could hinder the software adoption by EFL teachers as one of the methods used in their day-to-day combat against plagiarism. This could be attributed to the fact that even the free version could perform effective functions of determining originality in one document to many documents, a group of documents to another group, and multiple documents against each other within the local database.

It also suggests that the free version of the Plagiarism Checker X could be adopted including in but not limited to EFL classroom in order to prevent plagiarism. The adoption of paid versions could allow the teachers to check similarities between students’ papers with millions online published documents. This study adds to a growing body of literature on usability studies and the use of similarity checking software for preventing plagiarism.

REFERENCES


