The effectiveness of plagiarism detection software as a learning tool in academic writing education

Brad Stappenbelt
University of Wollongong, Australia

Chris Rowles
University of Western Australia

Abstract Plagiarism detection software (or more accurately, text-matching software) is commonly employed in a punitive capacity, detecting plagiarism after assignment submission. As an alternative to this approach, online plagiarism detection software was adopted as a learning tool for students instead. A trial was conducted in the foundation unit of the professional development component of the engineering degree at the University of Western Australia. Prior to the use of plagiarism detection software as a learning tool, efforts to instruct students regarding proper referencing and paraphrasing did not result in commensurate decreases in the levels of plagiarism detected. Many student assignments submitted displayed at the very least careless source acknowledgement. As part of the trial, students were given individual access to the software to self-assess their work as often as required prior to submission. The plagiarism detection algorithm assignment-originality statistics across three substantial written assignments throughout semester revealed continual and substantial improvement in student ability to avoid plagiarising. Through the use of this software, students were facilitated to learn how to properly acknowledge sources and improve their paraphrasing. This was accompanied by a dramatic decrease in the reportable incidence rates of plagiarism. Student perception of the use of plagiarism detection software in this capacity was also very positive.

Key Ideas

- Plagiarism detection software was adopted as a learning tool rather than as a plagiarism policy enforcement mechanism.

- The approach encouraged more experiential learning.

- The approach relieved some of the burden for teaching staff checking student work prior to submission.

- The approach assisted in building a community of academic integrity. Adopting plagiarism detection software as a learning tool, the educator's role was seen more as assisting writing skill development rather than policing plagiarism.

- There was a substantial 79% decrease in assignment first-draft mean level of plagiarism from the first to the second written assignment.

- There were no cases of plagiarism detected in the final assignment across approximately 620 students.

- Most students strongly agreed that access to the online plagiarism detection tool had been useful in their report preparations.

- Most students strongly agreed that the use of the online plagiarism detection tool had improved their ability to avoid plagiarising.

Discussion Question 1 What, if any, is the role of plagiarism detection software in developing and nurturing a community of academic integrity?

Discussion Question 2 Does the use of detection software facilitate the educational objective of transferring to students a sense of ethics and morality regarding plagiarism?

Introduction

Plagiarism, broadly defined as "passing off someone else's work, whether intentionally or unintentionally, as your own for your own benefit" (Carroll, 2002) is on the increase in higher education. The growth in information technology and accessibility has provided much material to fuel the observed increase in the incidence of plagiarism (Childs, 2001; McCabe, 2001; Maslen, 2003; Furedi, 2003). Tedford (2003) reported that more than 50% of high school students admit to engaging in plagiarising using the internet as their source. These are of course the students who will form our higher education cohort in subsequent years. The behavioural issues associated with students' plagiarising are complex and have been examined in numerous prior studies such as those described in McGowan (2005), Marsden, Carroll and Neill (2005) and Park (2003). A strong correlation has been demonstrated between the severity of academic dishonesty of students and unethical behaviour once they enter the workforce (Nonis & Swift, 2001). It is essential therefore, in producing future leaders in the community, that initiatives to foster academic integrity feature strongly in higher education institutions.

The use of plagiarism detection software in higher education was first notably implemented in 2001 at the University of Virginia (Tedford, 2003). In this well publicised case, a Physics Professor developed custom code to check 1500 student papers from the preceding three years. As a result of these checks, a number of students were investigated on plagiarism related academic dishonesty charges. More importantly however, the case served to highlight the lack of available information regarding the prevalence of this form of cheating in higher education and the minimal incorporation of plagiarism detection mechanisms in academic policy enforcement. Following this case, commercial plagiarism detection packages have increased rapidly in number and popularity. The commercial package *Turnitin* in particular has been adopted in a large number of higher education institutions and continues to be one of the preferred plagiarism detection alternatives available (Royce, 2003). The commonly adopted term 'plagiarism detection' software will be used to refer to the software throughout this paper. It must be noted however that the algorithm does little more than match text and would therefore be more accurately described by the term 'text-matching' software.

Despite the widespread tendency to place unquestioning trust in the results of online plagiarism detection algorithms, there are some significant inherent limitations of the methods employed (Martin, 2005). Firstly, an important distinction to make is that plagiarism detection software does not actually detect plagiarism. Rather, the software detects matching phrases (Royce, 2003). Some matches of student assignments with existing work should always be expected and accepted especially in fields of study with necessarily limited vocabulary. As a consequence, all cases suspected of plagiarism should always be checked by the instructor (Royce, 2003). Secondly, no search engine is capable of searching all available online material (Royce, 2003). Even the best search engines available search only a small portion of the internet. This is evidenced by the observation that different search engines return different hits with identical search terms. It is also impossible of course for the search engines to find matches with written material that has not yet been digitised. Text matching is also hindered by translation from different languages (Royce, 2003). Another potential limitation of plagiarism detection software is that students may simply learn to modify sentences or key words within a passage sufficiently so that they are not matchable with the source material (Royce, 2003; Martin, 2005). It is much more

difficult to detect and prove that ideas have been plagiarised than to do this for the passages and sentences that convey these ideas.

Methodology

The present academic integrity initiative in the Faculty of Engineering, Computing and Mathematics at the University of Western Australia, included the provision of online plagiarism detection software for student and staff use in evaluating written work. The commercial product *Turnitin* was employed for this purpose. This plagiarism detection software produces originality reports by comparing the submitted written material to existing text in the *Turnitin* database, online texts and journals and information from the internet (Frazer, Allan & Roberts, 2004).

The use of the online plagiarism detection software *Turnitin* was trialled in a core first year engineering unit. The unit involved in this initiative forms the foundation for the professional development component of the engineering degree. To succeed in this component of the degree a high level of written communication ability is required. Despite efforts to instruct students regarding proper referencing and paraphrasing in previous years, many students continued to submit written assignments that contain significant amounts of plagiarised material. Before the aid of plagiarism detection tools, approximately twenty severe cases of plagiarism were detected annually in this unit. The number of suspicious assignments that were never investigated was far greater, with an even larger number displaying at the very least careless source acknowledgement.

Plagiarism detection software is commonly employed in a punitive capacity, detecting plagiarism after assignment submission. The study by Martin (2005) for example demonstrated the long term benefits of using *Turnitin* in such a punitive capacity to significantly lower the incidence of plagiarism in a higher education setting. In the present initiative however, students were given individual access to the software to self-assess their work as often as required prior to submission as suggested by Baggaley and Spencer (2005) in their case study. To facilitate the adoption of the software, several lectures were conducted regarding proper source acknowledgement, referencing, citation and the use of *Turnitin*. The unit teaching staff, in particular the tutors were also available throughout semester to assist students in reaching the writing standards required.

Three substantial written assignments were set throughout semester. Students submitted drafts of the assignments to *Turnitin.* A three level screening process was used to examine the level of plagiarism reported by *Turnitin* through the text matches in the originality reports. Before class, students would identify sections of the *Turnitin* reports they believed to be incorrectly labelled by the software as being plagiarised. The tutors would then examine the written work and the accompanying originality report during in-class tutorial exercises. Students requiring additional assistance had their originality reports and written work further scrutinised by one of the four learning, language and research skills staff members teaching within the unit. Verified originality report statistics were collected for the first draft and final submission of each written assignment.

Results and Discussion

Student perception of the *Turnitin* software as a learning tool was generally very favourable with only a small number of students voicing concerns. The concerns raised by students were primarily related to technical software issues. There were no reported cases where students refused to use the software on the grounds of lack of fairness of the process or intrusion of their privacy as has been reported when the software is used solely in a punitive role (Tedford, 2003). The general acceptance of *Turnitin* is similar to the findings reported by Dahl (2007), where most students using the software were supportive of its adoption.

Using the UWA student perceptions of teaching (SPOT) survey tool, with response ratings from 1 (strongly disagree) to 5 (strongly agree), most students in the present initiative strongly agreed that access to the online plagiarism detection tool had been useful in their report preparations. Most students also strongly agreed that the use of the online plagiarism detection tool had improved their ability to avoid plagiarising. Mean student ratings of the usefulness of the online plagiarism detection tool in report preparation and its influence on their ability to avoid plagiarising were 4.2 and 4.1 respectively. The *Turnitin* report statistics (in particular the overall similarity index) for the three reports completed within the unit concur. These demonstrated significant and consistent improvement throughout semester in student abilities to properly paraphrase and reference material.

The *Turnitin* overall similarity indices (i.e. percentage of material matching internet sources, publications or student papers) for the first draft of the three written assignments set within the unit are presented in Table 1. In determining these percentages, material contained within quotation marks and reference lists were not included. Text matches of three words or less were also ignored. The Turnitin statistics show a substantial 69% decrease in assignment first-draft mean level of plagiarism from the first to the second written assignment. In the final (third) assignment submissions, similarity indexes for all 618 students were 24% or less.

It is unlikely that the positive results reported in Table 1 are the consequence of the majority of students simply learning to modify sentences or key words within a passage sufficiently so that they were not matchable with the source material. This practise has been noted within the present trial, but the number of students identified as using *Turnitin* in this manner was very small. The intent to deliberately engage in deceptive practice is generally not the governing motivator for students as discussed in the work by Deckert (1993). Most students are genuinely interested in learning. The mechanistic application of citation and referencing rules, involving trial and error phases using the *Turnitin* software, may also be a necessary initial learning stage on the path to competent academic writing.

Assignment	Overall Similarity Index			
	0-24%	25-49%	50-74%	75-100%
1	168	285	120	32
2	431	121	36	16
3	448	135	29	-

Table 1 – Overall similarity index for the first draft of three written assignments (n=618).

The success of the approach in reducing plagiarism and the associated education of students regarding proper citation and paraphrasing are believed to be due to several factors. Firstly, the approach encouraged more experiential learning. Rather than being involved primarily in passive instruction, students were actively engaged in, for example, repeated attempts to improve their paraphrasing using *Turnitin*. Using the software, students received frequent feedback regarding the originality of their written work and whether sources had been properly acknowledged. This feedback supplemented the necessarily limited feedback teaching staff of such a large unit of study could provide. The use of *Turnitin* consequently also had the desirable effect of relieving some of the burden for teaching staff in checking student work prior to submission. Finally, although the software was not used in a punitive capacity, it retained the effect of highlighting plagiarisms transgressions. Few students dared submit a piece of written work for assessment that they knew to have an overall similarity index above 24%.

Almost 19% of the students in the foundation unit were international enrolments (n=98). This group was responsible for the majority of the worst cases of plagiarism evident in the first draft of the three written assignments (see Table 2). This is consistent with qualitative observations from previous years. The international student group was also the most resistant to educational efforts to improve their writing. Contributing factors are likely to be the difficulty of working in another language, lack of writing instruction at high school level and the high pressure to succeed in light of the level of investment in their tertiary education (Song-Turner, 2008). Most importantly however, there appears to be a cultural misalignment in the perception of what constitutes plagiarism and why it is wrong to use sources without suitable acknowledgement. Although many of the international students consulted regarding their plagiarism openly admitted to the insertion of other people's work without acknowledgement, they genuinely did not appear to understand that they had committed an offence.

Assignment	Overall Similarity Index			
	0-24%	25-49%	50-74%	75-100%
1	31	29	23	15
2	47	32	7	11
3	45	41	12	-

Table 2 – International student overall similarity index for the first draft of three written assignments (n=98).

Adopting plagiarism detection software as a learning tool, the educator's role was seen more as assisting writing skill development rather than policing plagiarism. Because the students were essentially self-correcting any writing containing plagiarised material, the approach appeared to assist in building a community of academic integrity. The authors observed no signs that a culture of suspicion, resulting from plagiarism detection software adoption, was developing as described by Tedford (2003). Instead, many of the teaching staff reported positively on the experience. As discussed by Williams (2007), where a sense of betrayal and consequent distrust of students was described as a common response by educators when confronted with cases of plagiarism, the educational framework adopted in the present initiative served as an opportunity for better teaching.

Conclusion

Use of the *Turnitin* plagiarism detection software in an educational capacity appeared to be very successful in teaching students proper paraphrasing and source acknowledgement. As a consequence, it was also very effective in decreasing the incidence rate of plagiarism in the foundation unit of the professional development component of the engineering degree. Student perception of the usefulness and effectiveness of the use of the software was also very positive. Generally, due to the use of *Turnitin* and the accompanying focus this necessarily placed on the associated issues, students developed a sense that avoiding plagiarism is important. From this perspective, *Turnitin* appeared to be a successful learning tool in academic writing education.

Although compliance with the requirement for adequate source acknowledgement and paraphrasing of material was good, it was noted however that students were not automatically endowed with a moral sense regarding academic integrity. The mechanistic manner in which a software tool such as *Turnitin* encourages students to address the problem of plagiarism in their written work potentially limits the tool's usefulness in developing such an ethical perspective. As stated by Emerson, Rees and MacKay (2005) and Murray (2006), there is a need to provide students with a holistic perspective on the process of acknowledging sources not just the mechanics related to the conventions of citation. The authors agree that the adoption of online plagiarism detection must be accompanied by a commensurate educational effort targeted at improving student understanding of the reasons for avoiding plagiarism if a genuine community of academic integrity is to be encouraged.

Acknowledgements

References

Baggaley, J. & Spencer, B. (2005)'The mind of a plagiarist', Learning, Media and Technology, 30:1, pp. 55-62.

Carroll, J. 2002, 'A handbook for deterring plagiarism in higher education', Oxford Centre for Staff and Learning Development, Oxford.

Childs, S. 2001, 'Editorial: Web dependency—plagiarism', Health Information on the Internet, The Royal Society of Medicine Press Ltd. No. 21, June, pp. 1–2.

Dahl, S., 2007, 'Turnitin: The student perspective on using plagiarism', Active Learning in Higher Education v. 8 no. 2, pp. 173-191.

Deckert, G., D. 1993, 'Perspectives on plagiarism from ESL students in Hong Kong', Journal of Second Language Writing, Vol. 2, No. 2, pp 131-148.

Emerson L., Rees M. & MacKay B. 2005, 'Scaffolding academic integrity : creating a learning context for teaching referencing skills', Journal of University Teaching and Learning Practice, v.2 n.3a, pp.12-24.

Page 6 of 7

4th Asia Pacific Conference on Educational Integrity (4APCEI) 28–30 September 2009
University of Wollongong NSW Australia
Refereed Paper

Frazer, J., Allan, G. & Roberts, R. 2004, 'Implementing Turnitin/Originality Verification software at RMIT University in the context of Academic Integrity', Available online: http://www.rmit.edu.au/academicintegrity.

Furedi, F. 2003, 'Shortcut to success', Times Higher Education Supplement, July 25, p. 16.

Marsden, H., Carroll, M. & Neill, J. 2005, 'Who cheats at University? A self-report study of dishonest academic behaviours in a sample of Australian university students', Australian Journal of Psychology, 57(1), pp. 1-10.

Martin, D. F. 2005, 'Plagiarism and Technology: A Tool for Coping with Plagiarism', Journal of Education for Business, v. 80 no. 3, pp. 149-152.

Maslen, G. 2003, '80% admit to cheating', Times Higher Education Supplement, January 24, p. 17.

McCabe, D. 2001, 'Student cheating in American high schools', Educause (online), Available from: http://www.academicintegrity.org (Accessed: June 2008).

McGowan, U. 2005, 'Does educational integrity mean teaching students NOT to 'use their own words'?', International Journal for Educational Integrity, Vol. 1, No. 1, Available from: http://www.ojs.unisa.edu.au/index.php/IJEI/article/view/16/6 (Accessed: July 2009).

Murray, W. 2006, 'The plagiarism phenomenon', E.learning Age, pp. 22-24. Retrieved August 10, 2009, from ABI/INFORM Global.

Nonis, S. & Swift, C. 2001, 'An examination of the relationship between academic dishonesty and workplace dishonesty: A multicampus investigation', Journal of Education for Business, 75 (6), pp. 321-330.

Park, C. 2003, 'In other (people's) words: plagiarism by university students – literature and lessons', Assessment & Evaluation in Higher Education, Vol. 28, No. 5, pp. 471-488.

Royce, J. 2003, 'Has Turnitin.com Got It All Wrapped Up?', Teacher Librarian, v. 30 no. 4, pp. 26-30.

Song-Turner, H. 2008, 'Plagiarism: Academic dishonesty or blind spot of multicultural education?', Australian Universities' Review, vol. 50, no. 2, pp39-50.

Tedford, R. 2003, 'Plagiarism detection programs: A comparative evaluation', College & University Media Review, v. 9 no. 2, pp. 111-118.

Williams, B. T. 2007, 'Trust, betrayal, and authorship: Plagiarism and how we perceive students', Journal of Adolescent & Adult Literacy v. 51 no. 4, pp. 350-354.

Page 7 of 7

4th Asia Pacific Conference on Educational Integrity (4APCEI) 28–30 September 2009
University of Wollongong NSW Australia
Refereed Paper