

The multiple uses of iThenticate in doctoral education: Policing malpractice or improving research writing?

Alistair McCulloch, Monica B. Behrend, Felicity Anna Braithwaite
University of South Australia

This article provides a description and analysis of the way in which research degree students and their supervisors at one Australian university (the University of South Australia) use a popular online plagiarism-detection system, iThenticate. The study identifies how these two groups use iThenticate by analysing usage data together with data from an anonymous online survey conducted 12 months after the university took out a pilot subscription to the system. One hundred and nineteen students and 26 supervisors responded to the survey, representing 61% and 43% of the active users in each category. The survey found that the two groups of respondents used the system differently but that, while for both groups iThenticate's regulatory function in preventing plagiarism (whether international or accidental) was important, the system's potential educational function in improving research writing capability and publication was equally important. The study highlights the value of regarding the use of anti-plagiarism software so as to encourage a move away from a simple focus on its punitive regulatory dimension and towards its educational possibilities and suggests directions for future research on the relationship between this type of software and the ways scholars work with other people's texts to recreate meanings and develop original contributions.

Implications for practice or policy:

- Online plagiarism detection systems (such as iThenticate) can be used either negatively to police doctoral students' practice or positively to improve their research writing practice.
- Academic developers should promote a positive approach, aimed at improving research writing practice, as the preferable pedagogy in using online plagiarism-detection systems.

Keywords: plagiarism, academic integrity, research writing, anti-plagiarism software, doctoral education, online survey

Introduction

Academic and research integrity are matters of high importance in contemporary higher education and the expectations and requirements about them are built into both university and national academic frameworks and codes of good practice on doctoral education (Australian Council of Graduate Research, 2018; Quality Assurance Agency, 2018). Those involved in teaching and learning in the coursework sphere aim to ensure that their students develop appropriate ways of reading, note-taking and writing which adhere to disciplinary and professional standards. Those involved in research training or doctoral education also aim to ensure that their research trainees and doctoral candidates graduate as fully-fledged ethical researchers, can conduct independent research with a high degree of integrity, an important part of which involves rejecting plagiarism as a concept and avoiding it in practice. Writing in 2015, Ison went so far as to argue that:

Perhaps the most troubling cases of plagiarism are those that occur at doctoral level. Because so much of the doctorate degree revolves around and depends upon the performance and completion of the dissertation, misrepresentation of material in this critical document essentially makes the degree itself illegitimate. (p. 156)

This focus on plagiarism has at its centre the idea that ethical research conduct involves inter alia the appropriate attribution of ideas and ways of expressing them to their originators and to others who have already written about them and that learning to understand this and putting it into practice is a vital part of becoming a graduate researcher.

Although plagiarism is not a new phenomenon, there are suggestions that, with the development of digital resources and tools and the Internet, plagiarism is both easier and more commonplace than was hitherto the case. (Ison, as we note shortly, avers from this.) In tandem with this development, the digital age also provides software which can identify plagiarism thereby allowing those charged with promoting ethical academic behaviour to students to either educate or penalise those found to be involved in it. Among the software packages most commonly used in higher education to detect plagiarism are Turnitin (for coursework programs) and iThenticate (for research degrees). The two packages have been developed and are owned by the same company and, therefore, have similarities.

Although there is now a developing scholarly literature on Turnitin and the way it can be used as a formative tool in the development of academic writing skills for students on taught programs (Alharbi & Al-Hoorie, 2020; Cohen, 2010; Davis, 2007; Davis & Carroll, 2009; Penketh & Bowman, 2014), the same cannot be said for iThenticate where there is virtually no empirical evidence about the ways in which it can be used as an educational tool. The scholarly literature on iThenticate (summarised later in this paragraph) focuses very firmly on its regulatory and punitive roles rather than on its possible role as an educative tool. This mirrors the dominant discourse in the literature on plagiarism which positions the behaviour firmly as academic malpractice viewing it from a *rule ethical* rather than a *care ethical* approach (Vehviläinen et al., 2018). Although the former “is based on the notion of the universal application of a rule viewed as detached from the subject”, the latter “places the act in relation to the actor’s intentions and understandings ... and is based on the idea of virtues as guiding principles and of nurturing the students’ growth in these virtues” (p. 2). This focus on the rule, regulatory or punitive role of plagiarism-checking software can be seen in review articles (Awasthi, 2019) and also in articles discussing the prevalence of plagiarism in papers submitted to journals (Choi et al., 2016; El-Tahan, 2019; Higgins et al., 2016; Kalnins, 2015; J. H. Lee, 2014; W. Lee, 2015; Lykkesfeldt, 2016; Smart & Gaston, 2019; Zhang & Xiaoyan, 2012). The same bias can be seen in editorial communications from journal editors to their potential contributors and readers (Berquist, 2013; Driggers, 2017; McCuen & Govindaraju, 2015).

There is a paucity of literature on plagiarism-checking software and the research degree space. In 2015, Ison used Turnitin to examine 384 PhD dissertations from Canadian and United States of America (USA) universities to ascertain whether those completed prior to 1994 (i.e., before the Internet had become pervasive) were more prone to plagiarism than those completed since 2010. Although Ison found that “more than half of all analyzed dissertations contain evidence of plagiarism”, he concluded that “the mean similarity indices and frequencies of plagiarism occurrence discount the notion that the Internet is the instigator or cause of increased textual misuse” (2015, p. 163). Two earlier USA studies had examined the prevalence of plagiarism in science, technology, engineering, maths and medicine graduate students’ research proposals (Gillmore et al., 2010) and the patterns observed in such plagiarism (Vieyra et al., 2013), the latter of which moved a little away from the rule ethical approach by suggesting that “all students should receive formative (non-punitive) feedback when plagiarism is detected” (p. 47).

This article aims to extend the boundaries of what we know about doctoral education and the use of plagiarism detection tools from the perspectives of both research degree candidates and their supervisors. It offers an analysis of the way in which iThenticate was introduced into an Australian university, how it was used in that setting and the purposes for which it was used. Essentially, it is about how research degree candidates and their supervisors experience and use systems like iThenticate. The data on which the analysis draws was gathered during an internal evaluation exercise conducted by staff then situated in the university’s Teaching Innovation Unit as part of the decision-making process as to whether the university should continue to subscribe to iThenticate at the end of an initial trial period. We would like to make it clear that neither iParadigms (the corporate owner of iThenticate) nor any of that company’s component companies have had any input into this article or the evaluation exercise on which it draws and that the university paid and continues to pay the full cost of the service provided by iThenticate.

iThenticate

An overview

iThenticate is large-scale commercial professional plagiarism prevention software designed for researchers including higher degree by research (HDR) candidates and academic staff. (HDR candidates, to use the Australian parlance, are called by various names across the world including doctoral students, graduate

students, and research students. In recognition of the difference between the pursuit of a taught or coursework program and a research-based program, the University of South Australia (UniSA) has recently changed its HDR nomenclature from student to candidate and this is used throughout the article.) iThenticate enables a user (who might be the author, a supervisor or a journal editor) to upload a draft (e.g., proposals, thesis chapters, articles, conference papers or a full thesis) and then provides a report identifying direct matches and a percentage figure showing the Similarity Index, which is computed by comparing the uploaded writing to the system's very large database of textual sources. Where the results identify close similarity or possible plagiarism of some parts of the text, an author may review their writing and revise the document, a supervisor might provide feedback and advice on writing and avoiding charges of plagiarism or an editor-in-chief may decide to return a manuscript without sending it out to reviewers. iThenticate is designed specifically for researchers and the system identifies areas of potential plagiarism by checking the submitted piece of writing against a wide range of published, conference and grey literature. iThenticate is used by an increasing number of universities, and many international publishers and academic journals use it routinely to screen paper submissions for plagiarism before sending them out to referees. It is also the engine behind CrossCheck, a plagiarism-checking system used extensively by academic publishers across the world.

Although iThenticate is similar in function to Turnitin and is provided by the same company, it has three major differences to Turnitin. The first is that iThenticate has been developed specifically with researchers in mind rather than students on coursework programs. The second is that iThenticate draws on a more extensive database of sources against which it checks submissions, including 49 million journal articles, conference proceedings and books; 105 million online and offline subscription content and research titles from 30 leading aggregators, databases and content provider (including the ProQuest thesis and dissertation database); and a growing archive of more than 70 billion current and archived web pages (iThenticate, 1998–2020). The third difference is that, until they are deleted (by the user), iThenticate retains copies of documents submitted by a user only within the user's record. This is unlike Turnitin, which, unless the user explicitly makes an exemption, retains a copy of submitted drafts within its repository as a means of detecting any copying of essays and other coursework across students on taught programs. This means that revised versions of a paper can be submitted to iThenticate on multiple occasions as it is developed, revised and improved without being flagged as having similarities to an earlier version of itself. This facility reflects both the iterative nature and the developmental dimension of research writing.

Use of iThenticate at UniSA, 2016–2017

UniSA began using iThenticate in June 2016, and this article draws on data collected as part of an evaluation of the usage and users' experiences of the system during the following year. During this period, HDR candidates and academic staff used the system, with anecdotal evidence indicating its usefulness both in detecting potential plagiarism and in enabling candidates to rewrite affected parts of their text so as to avoid it. Centrally based staff working with HDR candidates in the development of their research writing abilities provided feedback suggesting iThenticate also functioned as a tool which contributed to that development. The primary purpose of the evaluation was to determine its value to both the university's HDR candidates and its academic staff so as to inform a decision about whether or not to continue to subscribe to the product. The evaluation was designed to allow it to address a number of issues which were focused on the care ethical rather than simply on the more common rule ethical approach to plagiarism discussed earlier (Vehviläinen et al., 2018). Put into question form, these were:

- (1) What are the usage patterns of iThenticate by research degree candidates and staff across the university?
- (2) How is iThenticate being used by research degree candidates and staff?
- (3) How easy is iThenticate to use?
- (4) Does iThenticate assist in developing a better understanding of research integrity matters?
- (5) How can the use of iThenticate be best advertised and maximised?
- (6) How adequate is the support for iThenticate?

Drawing on evidence from answers to the first, second and fourth of these questions, this article addresses both descriptive and technical issues such as numbers of users as well as the purposes for which iThenticate was being used and its role in furthering the formative ends of research education.

Methods

Two online survey instruments (one for staff and one for HDR candidates) were developed using SurveyMonkey following extensive consultation and piloting involving staff within UniSA. The instruments were structured in line with, and included questions designed to address, the six elements listed above. They were very similar in structure and wherever possible asked the same questions although sometimes this was not possible – where results impacted by this are presented, this is noted. Both surveys were opened on 31 May 2017, and separate invitations to complete the relevant survey were sent to all UniSA staff and HDR candidates who were registered as iThenticate account holders at that date. The emails contained a link to the relevant survey. A reminder email was sent on 8 June, followed by a final reminder on 16 June. A small number of new users were registered during the period during which the survey was open and these new users were included in the mailing list for each of the relevant reminder emails. The survey was closed on 19 June 2017.

Under the UniSA Human Research Ethics Policy, the development of evaluation data involving surveys of staff and students does not require a priori ethics clearance. In line with the policy, upon entering the survey instrument, respondents were thanked for their “help in completing this short questionnaire about your experience of the iThenticate system” before being informed that:

The questionnaire is completely anonymous and you are free to answer as many or as few questions you choose. Please note that data collected through this survey will be used to inform improvements at UniSA and could also be used in external publications and presentations. Individual responses will remain confidential and no individuals will be identified.

In addition to the two surveys, two focus groups were held, run by one of the HDR candidate representatives on the university’s Research Degrees Committee. Associate deans: research education (or equivalent) were contacted and either interviewed or asked for input by email. Two ad hoc interviews with PhD candidates were also undertaken by the HDR candidate representative, and a number of individual emails were sent to individuals who had made extensive use of iThenticate, only one of which generated a response. These discussions and exchanges helped inform our understanding of the survey data, but none of the data collected through the interviews or focus groups is directly included in this article.

Respondents and their patterns of iThenticate usage

Table 1 shows the number of registered users, the number of active users (a user was defined as active if they had submitted any writing for checking during their period of registration), the proportion of users who were active users and finally the response rate for each of the two groups of users as a percentage of the active users of iThenticate at the close of the survey.

Table 1
Number of staff and HDR candidates registered to use iThenticate, number of active users of iThenticate and percentage of respondents to the evaluation survey

	No. of registered users	No. of active users	Active users as % of registrations	No. of survey respondents	Respondents as % of active users
Staff	145	61	42%	26	43%
HDR candidates	451	195	43%	119	61%

Response rates varied across the institution with the then Division of Education, Arts and Social Sciences being significantly under-represented in terms of both staff and HDR candidate respondents. It was not possible to determine why this should be the case.

Of the relatively small number of staff responding to the survey, 19 (83%) of the 23 who answered this question were supervising a UniSA HDR candidate, two (9%) had previously done so and only two had no experience of supervision either at UniSA or elsewhere.

Table 1 shows that similar proportions of registered staff (42%) and HDR candidates (43%) were active users of the system, but that a significantly higher proportion of active candidate users (61%) responded to the survey than did active staff users (43%).

During the period from the point at which iThenticate was first made available across the UniSA research community through to Friday 16 June 2017, a total of 1,887 different pieces of writing had been submitted to iThenticate. Given that iThenticate counts a document as being equivalent to 25,000 words or fewer, and that anything over 25,000 words is counted as two or more documents depending on length, this represented 2,122 iThenticate documents. To avoid confusion, in this article the word “document” is reserved for its iThenticate meaning while individual submissions of pieces of writing made by staff or HDR candidates are referred to as submissions or pieces of writing. The overall pattern of submissions is shown in Table 2.

Table 2
Usage patterns for iThenticate, June 2016 – June 2017

No.	HDR candidates	Staff
Submissions	1,267	620
Document count	1,484	648

These high-level figures obscure a significant degree of variation in the number of submissions made by active users (see Figure 1). As can be seen, similar patterns of total usage are seen across both groups (staff and candidates) with the majority making one or two submissions and a small percentage making a significant number. Staff were more likely than candidates to have submitted only a single item (48% as opposed to 38%.) Heavy users (those with more than 40 submissions) were contacted individually by email and asked to provide more detail about their usage of iThenticate. Only one (a staff member) responded saying that they were using it primarily “to check the similarity index of manuscripts or manuscript drafts co-authored by me and my students prior to (re)submission”.

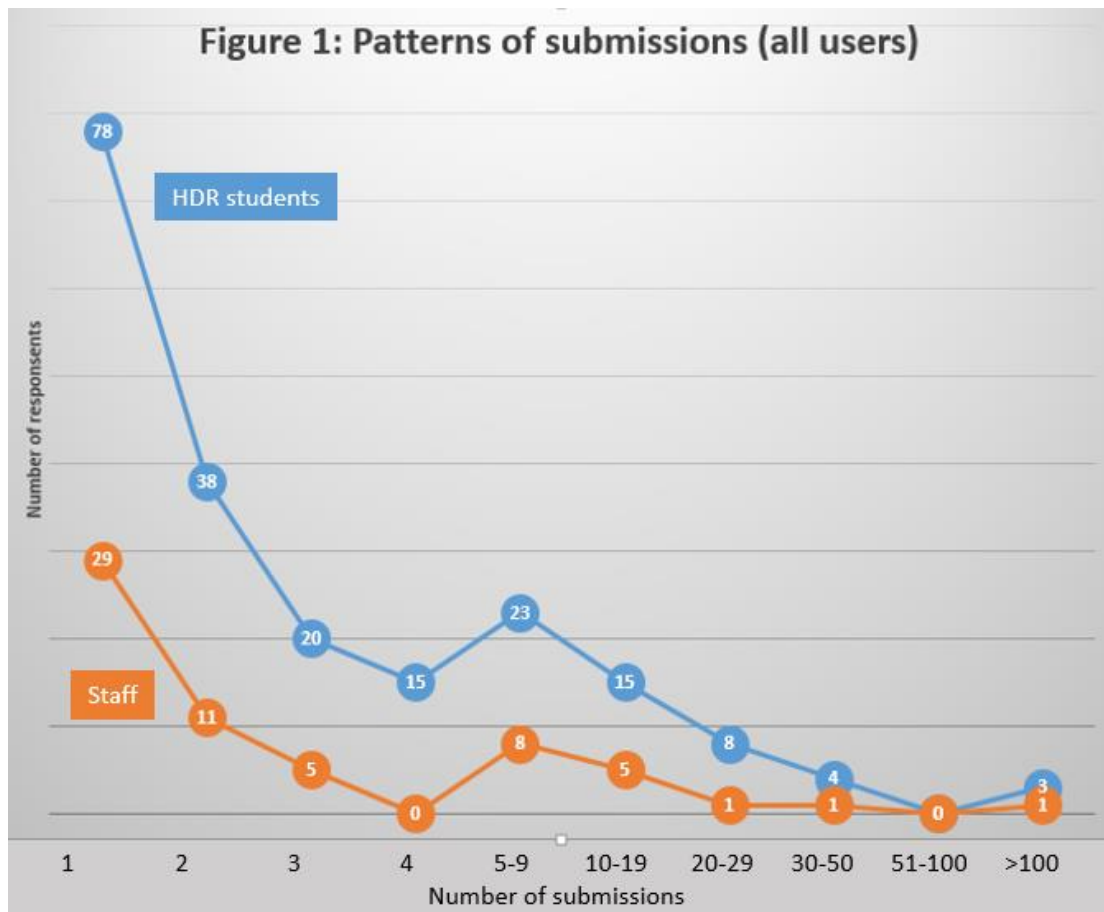


Figure 1. Patterns of submissions to iThenticate by academic staff and HDR candidates

Table 3 shows the number of submissions respondents reported having made. A small number of staff and HDR respondents (one and 13 respectively) reported not having used the system and they were directed past questions on the iThenticate user-experience and on to questions asking whether they felt confident about using the system. They were, nevertheless, asked why they had not submitted any writing to iThenticate and their answers related largely to their not perceiving a current need to do so, or in the case of candidates, because they had just commenced their HDR studies. The one staff respondent who reported not having used the system was generally negative about iThenticate saying they “could not see why we need it” and that they would not recommend it to either HDR candidates or other supervisors.

As can be seen in Table 3, both staff and HDR candidates reported similar patterns of usage as regards number of pieces of writing put through iThenticate, the modal usage being between 1 and 3 submissions.

Table 3
Number of individual submissions to iThenticate made by respondents

No.	HDR candidates		Staff	
	Number	Percentage	Number	Percentage
None	13	11%	1	4%
1–3	67	57%	14	58%
More than 3	37	32%	9	38%
Base	117	(100%)	24	(100%)

Differences between candidate and staff usage appear when respondents were asked about their practice regarding re-submission of the same piece of writing following an initial iThenticate check. Table 4 shows that staff are more likely to put a piece of writing through iThenticate once only, with none reporting resubmitting any item more than three times, whereas HDR candidates are most likely to submit and resubmit the same piece between one and three times, with almost one in five reporting that they had resubmitted the same document more than three times.

Table 4
Number of times respondents reported resubmitting the same piece of writing to iThenticate for a second or subsequent check following its initial submission

	HDR candidates		Staff	
	Number	Percentage	Number	Percentage
None	35	34%	15	65%
1–3	49	48%	8	35%
More than 3	19	18%	nil	nil
Base	103	(100%)	23	(100%)

What was iThenticate being used for?

Respondents were asked what types of writing they had submitted to iThenticate, and the responses are summarised in Table 5. Slightly different options were provided in the HDR candidate and staff surveys, and these are noted in the table.

Candidates and staff submitted drafts of thesis chapters in roughly a similar proportion, as was also the case for candidate’s final research proposals. Significant differences in use were seen with drafts (as opposed to final versions) of candidates’ research proposals, where almost six out of 10 candidates reported submitting this type of document in contrast to only one in eight staff, and drafts of writing intended for publication (through conferences and articles) which, perhaps not surprisingly, staff were more likely to have submitted. A variety of other types of writing were reported as having been submitted including drafts of full theses (two candidates and one staff member). The most common types of chapters to have been submitted by candidates were introductions and literature reviews, suggesting that those at an early stage of their studies may be more likely to adopt a new system than those who are further on in their studies and more set in their ways of writing and revising.

Table 5
Type of documents submitted to iThenticate

	HDR candidates No. (%)	Staff No. (%)
<i>Similar questions addressed to both candidates and staff</i>		
• Draft of 1 or more thesis chapters (candidate)	33 (32%)	
• Draft of 1 or more of a student's thesis chapters (staff)		6 (26%)
• Draft of my research proposal (candidate)	60 (59%)	
• Draft of a student's research proposal (staff)		3 (13%)
• Final research proposal (candidate)	33 (32%)	
• A student's final research proposal (staff)		5 (22%)
• Draft of conference paper or article (candidate)	38 (37%)	
• Draft of conference paper or article authored by a student (staff)		4 (18%)
<i>Staff-only questions</i>		
• A draft of a conference paper on which I was a main author	n/a	5 (22%)
• A draft of an article on which I was a main author	n/a	17 (74%)
Base	102	23

Note. The surveys used the language of student rather than candidate, and the new nomenclature had not yet been adopted and this is reflected in the language in this table.

iThenticate and the development of writing and understanding of research integrity matters

As noted above, plagiarism-checking software can be positioned as operating either primarily as a regulatory and punitive system or as also having a formative role to play. UniSA's evaluation was concerned to explore this second dimension of the system as well as the first and found evidence that iThenticate was being used as a formative tool useful not simply for identifying plagiarism, but also as an educative tool assisting individuals (and also those who work with them such as supervisors in the case of HDR candidates) to improve competence in research writing. As Table 6 shows, when asked "Did iThenticate help you prevent plagiarism?", 89% of HDR candidates said it had, with 64% of staff reporting that it had helped them prevent plagiarism in their candidates' writing. (About half of the staff responding also said that it had help them prevent plagiarism in their own writing.) It is important to note that this self-directed checking for plagiarism is not a simple case of *rule ethical* behaviour even though it may be driven by a desire not to infringe rules and thereby suffer sanctions, but also involves a significant element of self-directed *care ethical* behaviour. Assuming behaviour informs or at least facilitates learning, then using a plagiarism checking tool such as iThenticate implies some degree of formative learning.

Table 6
Answers to the question, "Did using iThenticate help you to prevent plagiarism?"

	HDR candidates (Only a single option could be selected)		Staff (More than one option could be selected)	
	Number	Percentage	Number	Percentage
Yes, in my students' writing	Not asked	n/a	11	64
Yes, in my own writing	Not asked	n/a	9	53
Yes	90	89	Not asked	n/a
No	11	11%	3	18%
Base	101	(100%)	17	(100%)

Note. The surveys used the language of student rather than candidate and this is reflected in the language in this table.

Moving beyond the simple detection of possible plagiarism, respondents were asked whether using iThenticate had helped them identify instances in their own writing where they had needed to make revisions (for whatever reason and not simply to prevent plagiarism) and, as Table 7 shows, over 80% of each group reported that it had.

Table 7
 Responses to the question, “Did using iThenticate help you to identify instances in your writing where you needed to make revisions?”

	HDR candidates		Staff	
	Number	Percentage (%)	Number	Percentage (%)
Yes	88	86%	18	82%
No	14	14%	4	18%
Base	102	(100%)	22	(100%)

The development of research writing skills is not simply a matter of avoiding plagiarism although, inevitably, the avoidance of plagiarism is a natural consequence of improved writing competence. Improving one’s writing skills is, of course, a key part of the process of developing as a researcher and a number of questions were asked about this to supplement our understanding of what, if any, impact iThenticate had on writing.

Respondents were asked whether using iThenticate had improved their writing (again, as opposed to simply helping them avoid plagiarism) and, as Table 8 shows, over 80% of both staff and candidates agreed it had, with about a third of each group “agreeing strongly” that it had and about half “agreeing”.

Table 8
 Responses to statement “Using iThenticate has improved my writing”

	HDR candidates		Staff	
	No.	Percentage	No.	Percentage
Strongly agree	30	29%	8	36%
Agree	55	54%	11	50%
Disagree	17	17%	1	5%
Strongly disagree	nil	nil	2	9%
Base	102	(100%)	22	(100%)

When staff respondents were asked whether using iThenticate had improved their candidates’ writing, about nine out of ten agreed that it had, positive responses being split almost half and half between those “strongly agreeing” and simply “agreeing” with the prompt statement. Two individuals (out of 15 respondents) said they disagreed strongly that it had done this (see Table 9).

Table 9
 Responses to statement “Using iThenticate has improved my student’s writing”

	Staff	
	No.	Percentage
Strongly agree	6	40%
Agree	7	47%
Disagree	nil	nil
Strongly disagree	2	13%
Base	15	(100%)

Paraphrasing is a key skill in the development of high-quality research writing and in the avoidance of inadvertent plagiarism. Taking the ideas and findings of other writers and putting them in one’s own words is a high-level skill that all HDR candidates should acquire during the course of their studies. Table 10 shows that high proportions (about 85% in each case) of both candidates and supervisors reporting that using iThenticate had improved their paraphrasing skills.

Table 10

Responses to the statement “Using iThenticate has improved my paraphrasing skills” (candidate statement) OR “Using iThenticate has improved my student’s paraphrasing skills” (staff statement)

	HDR candidates		Staff	
	Number	Percentage	Number	Percentage
Strongly agree	31	30%	5	33%
Agree	56	55%	8	53%
Disagree	15	15%	nil	nil
Strongly disagree	nil	nil	2	13%
Base	102	(101%)	15	(99%)

While the absolute numbers are small and care must be taken not to overstate the case, 11 (73%) of the 15 responding staff said that the use of iThenticate had enabled them to identify instances in their HDR candidates’ writing where they needed to make revisions. Further, about 80% of staff agreed that using the system had made them reconsider the way they themselves wrote. A similar proportion of HDR candidates also reported this same impact on their writing practice (see Table 11).

Table 11

Response to statement “The use of iThenticate has made me reconsider the way I write”

	HDR candidates		Staff	
	No.	Percentage	No.	Percentage
Strongly agree	29	29%	3	14%
Agree	49	49%	15	68%
Disagree	22	22%	2	9%
Strongly disagree	nil	nil	2	9%
Base	100	(100%)	22	(100%)

Finally in this regard, developing an understanding of and positive orientation towards research integrity is a key element in the development of a research trainee and about two-thirds of responding staff reported that using iThenticate had enabled them to educate their candidate(s) in this respect.

Recommendation as a test of value

One of the best tests of the value of a resource is whether those being asked to evaluate it would recommend it to others. Accordingly, staff were asked whether they would recommend iThenticate to HDR candidates they might supervise in the future. Almost all (88%) said they would with the following supportive comments being made in response to an open question asking for comments on the system. The first reflects a view of the system’s usefulness in securing future publications:

This software is an absolute necessity. Journal editors in my field will run submitted articles through this software to determine a matching score, they then use cut-off criteria and failing to meet this criteria can result in a desk rejection. Sometimes the match can be quite trivial but as authors we need to be aware of how the editors are evaluating the originality of submitted manuscripts.

The second reflects its educative role:

A useful educational tool, especially in learning how to move to substantial revision of text when working with research students.

Three staff said that they would not recommend it and when asked why that was the case, one said that there “is no obvious need at the moment” and a second that it:

Confuses the definition of plagiarism. The narrow re-wording needed to reduce an iThenticate match is not really paraphrasing, so this sends the message that paraphrasing is simply a thesaurus exercise rather than a re-expression of an idea.

In terms of recommending iThenticate to other HDR supervisors, 91% of staff respondents replied positively with only two saying they would not. When the two who would not recommend the system were asked why, one replied (again) that they had “not needed to use it so cannot see the point yet” and the other that there was a:

High number of false positives (up to 50% in one case) where the student incorporates their own published work into a thesis chapter ... I would recommend its use at the research proposal stage where supervisors maybe have not established the writing ability and discipline conventions are still being learned. By the time students are publishing or writing the thesis iThenticate is of limited usefulness.

HDR candidates were also asked whether they would recommend iThenticate to other candidates and almost all (97%) said that they would with only three responding negatively. Only one of those gave a reason for their negative answer and that was because they had “found the process difficult and stressful”. Among candidate responses were the following which pointed to the system’s usefulness in providing candidates with reassurance that they were maintaining high standards of research integrity:

iThenticate is a valuable resource to provide assurance that my work does not contain plagiarism.

The software provides tremendous support to the research student in maintaining academic integrity.

Excellent tool to assist and ensure I have cited correctly.

Other comments pointed to the usefulness of the system in providing formative feedback:

I thought it was a great way of making sure I was clearly articulating my own ideas and not those of others.

It is fantastic to have a useful, easy to use resource to assist with my learning and writing.

Discussion

This paper has delved into how both HDR candidates and academic staff members at one university use iThenticate. The paper presented data from a study designed to evaluate the way in which iThenticate was used within the doctoral education space. Data from the university’s iThenticate usage statistics and the two surveys with 119 HDR candidates and 26 staff members respectively showed clearly that the software platform was valued and useful both in terms of detection of plagiarism and also (and as importantly) as an educative process within itself. Collectively, candidates and staff valued iThenticate as a tool which could support publication, help them make substantial revisions to writing, and learn paraphrasing skills and how to better express their own ideas.

We note, perhaps unsurprisingly, the finding that candidate usage was dominated by the checking of research proposal documentation as opposed to the checking of thesis chapters or journal articles. This probably reflects greater awareness of iThenticate as a tool amongst more recent commencers than amongst candidates further along in their doctoral studies. With continuing access to iThenticate (on the basis of the evaluation from which the data used in this article was drawn, the university adopted an ongoing subscription to the system) and, through personal reflection on its utility, it is hoped that candidates will recognise that iThenticate is valuable for all types of research writing. Candidates agreed that, in addition to enabling the detection and prevention of plagiarism, iThenticate had a role in improving the overall quality of their own research writing, including by developing paraphrasing skills and requiring them to rethink their own research writing practice. This suggests that HDR candidates are interested in improving their own learning, which is an internal goal and one identified as an important reason for post-graduates to be less likely to be tempted to engage in plagiarism (Ives, 2020). Staff members predominantly used iThenticate for their own writing, rather than ascertaining the nature of plagiarism within the writing of their HDR candidate (although it did perform this latter function). Nevertheless, they acknowledged that iThenticate had the propensity to improve their candidates’ research writing and paraphrasing skills. Staff

were aware that journals are using software tools such as iThenticate to determine levels of plagiarism prior to an editor sending a manuscript out for review.

Despite the endorsement of the software provided by our data, further research is needed into specific and developing usage of iThenticate. This study was exploratory and limited by the relatively small number of respondents who were either academic staff (and thereby post-doctoral) or drawn from the Applied Social Science and Education academic areas. It also relied solely on survey data meaning that a more nuanced understanding of the sort allowed through qualitative interviews was not possible. Details of respondent demographics were not collected and, hence, analysis about usage differences based on English-language background was not possible. This is an area of possible future research as it has been claimed that this type of software mainly singles out non-native speakers of English or non-Western scholars (Introna & Haynes, 2011, as cited in Curno, 2015, p. 6). This survey was also not able to discern differences of practice between disciplinary practices, nor level of experience as research writers for HDR candidates or academics. As noted earlier, the indications are that most of the HDR candidate respondents were commencing candidates and mainly used iThenticate for checking their research proposals, something usually completed within the first 6–12 months of candidature. Further research could usefully engage with candidates who are moving towards or actively completing their theses.

The survey also ignored the issue of any pedagogical interventions or strategies that may have assisted candidates with their interpretation of and responses to iThenticate similarity reports. The relatively small number of academics and small amount of data about how supervisors of HDR candidates use iThenticate and discuss whether and how their candidates should or might engage in it would also provide a potentially fruitful avenue of further study. Usage data could be supplemented by interviews and the two used to explore practices related to the use of iThenticate; for example, how iThenticate reports are interpreted and used to improve research writing, or the types of discussions candidates have with supervisors and others who support the development of their research writing about iThenticate reports that best deepen understandings about citation practices. Interviews could also be used in future research to explore ways in which iThenticate is used in a formative way and its effectiveness in that respect. For example, it would be good to know if being aware that there was a need for better paraphrasing was enough to improve writing or whether something else needed. Did the awareness provided by iThenticate encourage a discussion between candidate and staff and, if so, who initiated such discussion? Did use of iThenticate encourage candidates to seek writing tutoring? These and other questions could usefully be explored in future studies.

Finally, the field of critical digital pedagogy has seen discussion around anti-plagiarism systems, surveillance technologies and issues of intellectual property, although most of the published work focuses on Turnitin because of that system's ubiquity across coursework programs and the retention by the system of any written work run through it (Amidon, 2016; Morris & Stommel, 2017; Vie, 2013). iThenticate differs in important ways and functionality from other anti-plagiarism systems and those using it are not simply demonstrating their understanding of prior works of scholarship and research but contributing to the body of work in supposedly original ways and in ways that will make their contribution publicly available. There are, therefore, different issues at play and future research could usefully focus on these differences and the way they play out as teachers and researchers learn to use anti-plagiarism tools more effectively as part of their pedagogical function rather than as part of their policing or regulatory function.

Conclusion

In conclusion, although the article reports on the use of only one commercially available text-detection system, iThenticate, we believe that the argument and the conclusions drawn are likely to apply to any digital plagiarism detection system as these systems differ largely in terms of technical detail rather than substance or purpose. The study highlights the value of the system in developing one aspect of research writing skills and in moving beyond the punitive regulatory dimension and into the educational sphere. We concur with Suchomel and Brandeys when they said that the “main goal of plagiarism detection systems is to improve the quality of textual works.” (2015, p. 18). While there are limitations (noted above) to the use of surveys in this area of research and evaluation, we believe that they can provide valuable data about research writers' experiences, about their use of text-detection software and also insights into knowledge and attitudes toward plagiarism as was the case, for example, in Memon and Mavrincac's (2020) survey study of early career researchers from mostly developing countries into their practices related to avoiding plagiarism. Given our finding that iThenticate is used by researchers and supervisors more in a formative

rather than a punitive spirit, we recommend that further research be conducted in association with both becoming and also more experienced researchers about the ways scholars work with other people's texts to recreate meanings and original interpretations relevant for their own specific writing purposes, whilst acknowledging the original ideas and the sources from which they are drawn.

Acknowledgments

The authors would like to thank Professor Pat Buckley, Dean of Graduate Studies at the University of South Australia, for her support during the project on which this article is based.

References

- Australian Council of Graduate Research. (2018). *Good practice framework for research training*. <https://www.acgr.edu.au/good-practice/best-practice/>
- Alharbi, M. A., & Al-Hoorie, A. H. (2020). Turnitin peer feedback: controversial vs. non-controversial essays. *International Journal of Educational Technology in Higher Education*, 17, 1–17. <https://doi.org/10.1186/s41239-020-00195-1>
- Amidon, T. (2016, September 8). (dis)Owning tech: Ensuring value and agency at the moment of interface. *Hybrid Teaching: Pedagogy, People, Politics*. <https://bit.ly/3s35Rr8>
- Awasthi, S. (2019). Plagiarism and academic misconduct A systematic review. *DESIDOC Journal of Library & Information Technology*, 39(2), 94–100. <https://doi.org/10.14429/djlit.39.2.13622>
- Berquist, T. H. (2013). Scientific integrity: Has CrossCheck made a difference? *American Journal of Roentgenology*, 201(4), 703. <https://doi.org/10.2214/AJR.13.11557>
- Choi, J., Park, S., & Oh, U. (2016). CrossCheck usage in a journal publication. *Science Editing*, 3(1), 26–32. <https://doi.org/10.6087/kcse.59>
- Cohen, J. (2010). Using Turnitin as a formative writing tool. *Journal of Learning Development in Higher Education*, 2. <https://doi.org/10.47408/jldhe.v0i2.22>
- Curno, M. J. (2015). Challenges to ethical publishing in a digital era. *Journal of Information, Communication and Ethics in Society*, 14(1), 4–15. <https://doi.org/10.1108/JICES-08-2015-0026>
- Davis, M. (2007). The role of Turnitin within the formative process of academic writing: A tool for learning and unlearning. *Brookes eJournal of Learning and Teaching*, 2(2). <https://bit.ly/3xsoIN5>
- Davis, M., & Carroll, J. (2009). Formative feedback within plagiarism education: Is there a role for text-matching software? *International Journal for Educational Integrity*, 5(2), 58–70. <https://doi.org/10.21913/IJEI.v5i2.614>
- Driggers, R. (2017). Ethics, plagiarism, and Crosscheck. *Applied Optics*, 56(16), ED5. <https://doi.org/10.1364/AO.56.000ED5>
- El-Tahan, M. R. (2019). Can the similarity index predict the causes of retractions in high-impact anesthesia journals? A bibliometric analysis. *Saudi Journal of Anaesthesia*, 13(5), 2. https://doi.org/10.4103/sja.SJA_709_18
- Gilmore, J., Strickland, D., Timmerman, B., Maher, M., & Feldon, D. (2010). Weeds in the flower garden: An exploration of plagiarism in graduate students' research proposals and its connection to enculturation, ESL, and contextual factors. *International Journal for Educational Integrity*, 6(1), 13–28. <https://doi.org/10.21913/IJEI.v6i1.673>
- Higgins, J. R., Lin, F. C., & Evans, J. P. (2016). Plagiarism in submitted manuscripts: incidence, characteristics and optimization of screening—case study in a major specialty medical journal. *Research Integrity and Peer Review*, 1(1), Article 13. <https://doi.org/10.1186/s41073-016-0021-8>
- Ison, D. C. (2015). The influence of the Internet on plagiarism among doctoral dissertations: An empirical study. *Journal of Academic Ethics*, 13(2), 151–166. <https://doi.org/10.1007/s10805-015-9233-7>
- iThenticate. (1998–2020). *Publish with confidence*. Retrieved August 9, 2021, from <http://www.ithenticate.com/>
- Ives, B. (2020). Your students are cheating more than you think they are. Why? *Educational Research: Theory and Practice*, 31(1), 46–53. <https://bit.ly/3CsRvEY>
- Kalnins, A. U., Halm, K., & Castillo, M. (2015). Screening for self-plagiarism in a subspecialty-versus-general imaging journal using iThenticate. *American Journal of Neuroradiology*, 36(6), 1034–1038. <https://doi.org/10.3174/ajnr.A4234>

- Lee, J. H. (2014). Analysis of CrossCheck data on two years' worth of papers submitted to archives of plastic surgery. *Archives of Plastic Surgery*, 41(5), 449–451. <https://doi.org/10.5999/aps.2014.41.5.449>
- Lee, W. (2015). CrossCheck data of manuscripts submitted to *Journal of the Korean Association of Oral and Maxillofacial Surgeons*. *Journal of the Korean Association of Oral and Maxillofacial Surgeons*, 41(3), 117–118. <https://doi.org/10.5125/jkaoms.2015.41.3.117>
- Lykkesfeldt, J. (2016). Strategies for using plagiarism software in the screening of incoming journal manuscripts: recommendations based on a recent literature survey. *Basic & Clinical Pharmacology & Toxicology*, 119(2), 161–164. <https://doi.org/10.1111/bcpt.12568>
- McCuen, R. H., & Govindaraju, R. S. (2015). Broad perspective on plagiarism. *Journal of Hydrologic Engineering*, 20(5). <https://bit.ly/2X5SH12>
- Memon, A. R., & Mavrinac, M. (2020). Knowledge, attitudes, and practices of plagiarism as reported by participants completing the AuthorAID MOOC on Research Writing, *Science and Engineering Ethics*, 26, 1067–1088. <https://doi.org/10.1007/s11948-020-00198-1>
- Morris, S. M., & Stommel, J. (2017, June 15). A guide for resisting Edtech: The case against Turnitin. *Hybrid Teaching: Pedagogy, People, Politics*. <https://bit.ly/3IGWXOz>
- Penketh, C., & Beaumont, C. (2014). 'Turnitin said it wasn't happy': Can the regulatory discourse of plagiarism detection operate as a change artefact for writing development? *Innovations in Education and Teaching International*, 51(1), 95–104. <https://doi.org/10.1080/14703297.2013.796721>
- Quality Assurance Agency. (2018). *The revised UK quality code for higher education*. <https://bit.ly/37uUizh>
- Smart, P., & Gaston, T. (2019). How prevalent are plagiarized submissions? Global survey of editors. *Learned Publishing*, 32(1), 47–56. <https://doi.org/10.1002/leap.1218>
- Suchomel, Š., & Brandejs, M. (2015). Source retrieval for plagiarism detection. *Journal of Advances in Information Technology*, 6(1), 18–26. <https://doi.org/10.12720/jait.6.1.18-26>
- Vie, S. (2013). A pedagogy of resistance toward plagiarism detection technologies. *Computers and Composition*, 30(1), 3–15. <https://doi.org/10.1016/j.compcom.2013.01.002>
- Vieyra, M., Strickland, D., & Timmerman, B. (2013). Patterns in plagiarism and patchwriting in science and engineering graduate students' research proposals. *International Journal for Educational Integrity*, 9(1), 35–49. <https://doi.org/10.21913/IJEI.v9i1.846>
- Vehviläinen, S., Löfström, E., & Nevgi, A. (2018). Dealing with plagiarism in the academic community: emotional engagement and moral distress. *Higher Education*, 75(1), 1–18. <https://doi.org/10.1007/s10734-017-0112-6>
- Zhang, Y., & Jia, X., 2012. A survey on the use of CrossCheck for detecting plagiarism in journal articles. *Learned Publishing*, 25(4), 292–307. <https://doi.org/10.1087/20120408>

Corresponding author: Alistair McCulloch, alistair.mcculloch@unisa.edu.au

Copyright: Articles published in the *Australasian Journal of Educational Technology* (AJET) are available under Creative Commons Attribution Non-Commercial No Derivatives Licence (CC BY-NC-ND 4.0). Authors retain copyright in their work and grant AJET right of first publication under CC BY-NC-ND 4.0.

Please cite as: McCulloch, A., Behrend, M. B., & Braithwaite, F. A. (2022). The multiple uses of iThenticate in doctoral education: Policing malpractice or improving research writing? *Australasian Journal of Educational Technology*, 38(1), 20-32. <https://doi.org/10.14742/ajet.7100>