LTC15D188

Title:	**Future use of Text-Matching Systems (TMS) in the Investigation of Plagiarism & Collusion Cases – A Comparative Evaluation of SafeAssign and Turnitin**
Author:	Adam Longcroft (ADTP), with support from School Plagiarism Officers, CTEL and LTS staff
Date:	July 2016
Circulation:	LTC – 6 July 2016
Agenda:	LTC15A007
Version:	Final version
Status:	Open

Abbreviations

ORs: Originality Reports (reports on a secript generated by a TMS which flag 'matched' sections of text in students' scripts with original sources)

P & C   Plagiarism & Collusion

TMS   Text Matching Software

CTEL   Centre for Technology Enhanced Learning, UEA

Issues

The background

During 2014/15 University LTC endorsed a shift from the use of Turnitin to SafeAssign as the University's approved Text-Matching Software (TMS). Turnitin is a subscription-based service, whereas SafeAssign is one of the in-built 'tools' within Blackboard VLE. During the current academic year (2015/16) it has become apparent that the performance of SafeAssign has fallen below the standard expected, in terms of its ability to highlight matches with material in online journals and websites etc, and in terms of the usability of the Originality Reports as a form of evidence which informs the investigation of suspected cases of P & C by markers and Plagiarism Officers. The PVC Academic subsequently tasked the ADTP with carry-out a comprehensive comparative evaluation of the two TMS systems, with the support of colleagues in the CTEL, LTS staff (Team leaders) and School Plagiarism Officers/Deputy Plagiarism Officers. The evaluation was carried out between May and June 2016, and Plagiarism Officers were required to submit their reports to the ADTP by 10am on 22 June 2016.

The ADTP sent a detailed explanation of the process by which the evaluation would be conducted to all Plagiarism Officers on 6.5.16 (See **Appendix 1**). In addition to clarifying the nature of the evaluation process, the e mail also set out the criteria for establishing the relative merits of the two TMS systems, and provided POs with a standard proforma (template) they should utilise for drafting their summary reports to University LTC (See **Appendix 2**). Copies of the Summary Reports produced by School Plagiarism Officers on the comparative evaluation of the two TMS systems are attached to this paper (See **Appendix 3**).

Reports received from Plagiarism Officers by the 22 June deadline, are as follows:

Received	Not received
MED	EDU (new PO in post)
DEV	PPL
LDC	AMA
HUM	HIS
BIO	MTH (new PO in post)

HSC	PHA
NBS	ECO
INTO	CMP
PSY	CHE
SWK	
LAW	
ENV	

Drawing a distinction between Investigation of cases, and batch screening.
TMS systems are used in Universities in two key ways:
1) For **batch screening** – where the scripts of an entire module cohort are submitted to a TMS as a 'check' prior to or following the marking process.
2) For **Investigative Purposes** – where a concern flagged by a marker is then investigated by a Plagiarism Officer (or the equivalent).

This paper sets out the results of the comparative analysis for University LTC and focuses on the use of TMS for **Investigative Purposes** – i.e. those instances where a marker has identified a concern and flagged this for the School Plagiarism Officer to carry out a further investigation in accordance with the University's P & C Policy.
The issue of which TMS system should be used for batch screening is a separate issue, and is likely to be the subject of a separate report for a future meeting of University LTC.

Resource Implications
1) **Costs** - If University LTC endorses the key recommendation set out below, then UEA will need to renew its subscription to Turnitin during 2016/17 (and beyond). This is likely to be the focus of a negotiation between CTEL colleagues and Turnitin. University LTC needs to be aware, however, that UEA will have little or no control over future subscription fees. If LTC decides to retain SafeAssign this will be a zero cost option.
2) **Technical implications** - If Turnitin is only used for Investigative Purposes, there are no significant technical implications, especially with regard to the online marking project which is already well advanced and which uses Blackboard as the submission point and marking tool for student scripts. For example, there would be no need to create a feed from BB to Turnitin etc. In the past, UEA has utilised Turnitin only for Investigative Purposes – so endorsing the recommendation will simply mean a reversion to past practice, and practice that Schools will be familiar with.
3) **Training implications** - Plagiarism Officers have all been provided with training on the use of Turnitin, so there will be few training implications for the University if Turnitin is only used for Investigative Purposes. However, POs who are newly appointed to the role in 2016/17 will require training on Turnitin.
4) **Staff time involved in apply the P & C Policy** - It has become increasingly clear that the choice of TMS systems has a considerable impact on the amount of staff time that is required to investigate P & C cases – especially those which are more complex or which involve larger submissions (e.g. projects and dissertations). PO's reports make it clear that the amount of time required to investigate cases and to prepare for Plagiarism Hearings is greatly increased because of the limited functionality of SafeAssign and the unhelpful way in which ORs present evidence which might be used to inform an investigation and/or a plagiarism meeting with the student.

Risk Implications
- If the recommendation is endorsed and Turnitin is only used for Investigative Purposes, there are few significant risk implications for university practice or technical infrastructure since this would involve simply a reversion to past practice.
- However, University LTC members should be aware that the choice of TMS systems is important in terms of providing an additional safeguard with regard to the integrity of the awards the University confers on its students. The more effective the TMS system in

identifying 'matches' (with material on websites, in journals, and in previously submitted student coursework), the more likely it is that instances of either unwitting or deliberate academic malpractice will be identified and penalised in accordance with the University's P & C Policy. Plagiarism Officers feel strongly that the integrity of awards may be threatened if the University's approved TMS system fails to perform to a satisfactory standard.

Equality and Diversity
The revised policy is extremely unlikely to have any negative consequences for students or staff with protected characteristics.

Timing of decisions
Approval by LTC the recommendation will drive a shift to the use of Turnitin as the University approved TMS for investigative purposes from 2016/17 onwards – i.e. the investigation of suspected cases of P & C where a marker flags a concern for the School PO to investigate further.

Recommendation
Having considered the very detailed and comprehensive reports provided by POs on the comparative evaluation of SafeAssign/Turnitin, the ADTP's recommendation to LTC is that the University should adopt **Turnitin** as its preferred TMS for **investigative purposes** from 2016/17 onwards.

Summative Report from the Academic Director for Taught Programmes (ADTP) on the Comparative Evaluation of SafeAssign and Turnitin TMS Systems.

CRITERIA
The evaluation focuses on the following criteria for establishing the relative merits/strengths of SafeAssign and Turnitin:
Speed
How long does it take to run-off Originality Reports using the two TMS systems?
It is clear from the reports submitted by POs that the speed with which the two TMS systems generate ORs is comparable, and in both cases can vary according to the time of day the script is submitted to the system, and the size/length of the script. The two TMS systems appear, therefore, to perform to a similar standard in terms of the time taken to run-off ORs. **Comments from POs:** *"My recollection is that Turnitin can take slightly longer in the afternoon to upload but I have never had a problem or frustration with either system in this respect"* (HSC). *"The time it takes to generate originality reports (OR) using SA or Turnitin is similar. Depending on time of day and the length of the submitted piece of work it might take between 1min and 1hr"* (LAW). *"Both have a similar speed. I don't need the software to analyse large batches of information at once, so the speed of both to analyse scripts I upload is fine"* (SWK). *"I've found no important difference in speed between the two"* (INTO). *"Didn't notice a significant difference in the upload time. SA is slightly easier to upload (as it doesn't require logging the name of the student and the name of the paper)"* (NBS). *"At quiet times, both perform at a similar speed. Occasionally there is some delay while Turnitin generates its report, but this has never been a problem for me as I usually submit a script and then return to the completed report later"* (NBS). *"I have found both tools reasonably fast to use"* (HSC). *"From experience, Turnitin can be a bit slow at times of peak demand"* (MED).

"Comparable. Turnitin seems to take slightly longer during periods of high demand, but this is vastly outweighed by the usefulness and customizability of the reports" (DEV).
"Speed of working of either software is not perceived to be an issue" (MED).
"From my experience, I don't see much difference between the two TMS systems" (LDC).
"Both Turnitin and SafeAssign ran promptly during my tests with each taking less than a minute to run a single essay of about 4000 words" (ENV)

Note for LTC Members: This is unlikely to provide a useful basis on which to base the decision as to which TMS system University LTC should approve.

Ease of use
How user-friendly is the TMS software and how easy is it to interpret the Originality Reports generated?

It is clear from the reports submitted by POs that there are serious concerns amongst the majority of POs about the 'Ease of Use' relating to SafeAssign. POs are almost unanimous in their view that Turnitin is much easier to utilise as a tool for investigating suspected cases of P & C, sharing evidence prior to Plagiarism meetings with students (as we are required to do within our P & C Policy), and using the OR during the Plagiarism Meeting itself to illustrate issues or query the use of sources etc. Many POs have made the point that use of SafeAssign adds considerably to the amount of time taken to investigate cases – sometimes multiplying the time required by a factor of 5 or more. POs have expressed a concern that continued use of SafeAssign could act as a serious disincentive for staff to flag concerns because of the time required to interpret SafeAssign ORs and to establish the extent of malpractice. It is also clear from the PO's reports that the raw OR 'scores' produced by SafeAssign are often entirely misleading with regard to the extent of actual plagiarism in a way which is much more problematic than with Turnitin. Certain sections of text can be 'filtered-out' by Turnitin (e.g. Bibliographies; fragmentary or small matches) but not in SafeAssign. These make the process of identifying the extent of actual plagiarism much more difficult, time-consuming and frustrating. It also makes communicating the basis on which concerns are based to the student much more difficult.

Comments from POs:
"Both Turnitin and SafeAssign are easy to initiate in terms of uploading a student's work and setting it going. However, the way that SafeAssign presents its outputs is much less clear and it more difficult to interpret the amount of plagiarism that has really occurred" (ENV).
"I find Turnitin the most easy to use… Managing the information, interpreting the information, the visual presentation of the information and the overall satisfaction with Turnitin's performance is far and above that of SafeAssign in ease of use" (HSC).
"Having no link between the highlighted sentence in the text and the juxtaposed suggested source at the end of the OR is a very careless omission on the part of the SafeAssign designers. This makes it very difficult to navigate the matches – both in investigation and when using in a meeting with a student" (MED).
"Functionality of SafeAssign is much less intuitive than Turnitin. Assignment has lost its formatting within the SafeAssign originality report, making it less clear to read and interpret, whereas this was not the case in Turnitin" (PSY).
"Interpretability of the scores produced in OR's are also an important aspect of ease of use. The 'similarity score' by Turnitin is presented as the proportion of the text which has been matched with another source. It seems to correspond with the volume of highlighted text. Neither the 'overall SafeAssign score' nor the 'sentence-matching score' produced by SafeAssign are interpretable – Blackboard gives conflicting information, none of which can yet be verified" (MED).
"SafeAssign is not user-friendly at all, especially once the originality report is shared with the marker (for further investigation) or with the student (for the plagiarism meeting). The ease of use is directly linked to detected matches. The matches that SA does pick up (and these are hardly any) are not highlighted in full…. Percentages are highly misleading and often wrong. Turnitin, however, does highlight the entire paragraph in colour which is off tremendous help in identifying the extent of the potential plagiarism at a glance" (LAW)

"The Turnitin software is easier to interpret. It is colour coded and this translates onto the pages which are then able to be downloaded and copied for the student to examine as well. This is not the same for SafeAssign" (SWK).

"I find both systems easy to use in terms of upload. Turnitin, however, allows for greater ease in presenting the results of detected plagiarism" (HUM).

"I did not find any difference in accessibility, speed or how user-friendly the applications were. I had no prior experience with either. I would safely say that based on this I would be happy to use either application" (NBS).

"Generally, I have found the speed of SafeAssign and Turnitin to be comparable in terms of actually generating the originality report, although the former (SafeAssign) is considerably more 'clunky' to use and it can take longer overall to process an assignment for this reason" (PSY).

"It is marginally more straightforward to generate a SafeAssign report, but this is mainly due to the fact that it has fewer useful options: e.g. the ability to filter out very short phrases, and exclude bibliographies and sections in quotation marks etc" (NBS).

"Turnitin is easier to use. It is very simple to understand and interpret the report…. It highlights the exact words which have been matched and links these to the source where they were found. The different matched sources are colour-coded, numbered, and listed in the order of largest percentage of matched text. SafeAssign does none of these things! Instead it highlights whole sentences which contain some matched text – it is not clear from the report which words have been matched. While the online SafeAssign report is in colour, it is not possible to download a colour version to show or even email to the student. The list of sources at the end is not numbered so can be very confusing where there are many different sources, and almost impossible to explain to a student in a meeting" (NBS).

"The percentage score generated by Turnitin is crude, but simple and consistent… The percentages generated in SafeAssign are far less clear" (NBS).

"I have found both systems reasonably easy to use, once I became familiar with them. The Turnitin reports are clearer and easier to read than the SafeAssign reports" (HSC).

"Two recent cases were dissertations. Both were brought to me by the marker, who had concerns that there may be plagiarism. In both cases, Turnitin was far more effective. It identified a larger percentage of matches. On occasion when I have used SafeAssign, a significant amount of additional material, not identified by SafeAssign, was identified by myself, via manual searching (very time consuming)" (HSC).

"The Turnitin reports are much easier to interpret. The different sources are colour coded. Verbatim matches are highlighted, so you can see which words were changed. The Turnitin report matches are logically numbered in the order they appear in the text. In contrast, SafeAssign highlights contiguous blocks of text, and reports a percentage match of dubious utility for the whole block, so you can't see exactly what has been changed" (MED).

"Turnitin is vastly more user-friendly, and the originality reports it generates are much easier to use…. Because Turnitin keeps the formatting the same, it is easier to compare the originally submitted version with the originality report" (DEV).

"Downloading an OR in SafAssign has to be done via a printing icon, and print to PDF – there is no other way of downloading an OR to store offline. I have not seen any instructions about this from BB" (MED).

"When it comes to interpreting the originality report, there is no doubt that the reports generated by Turnitin are overall more user-friendly, and, above all, easier to understand for the students when they are summoned to a hearing" (LDC).

Matches are all other student papers; appears to have missed large sections of plagiarised text in the introduction and failed to detect journal article sources detected by Turnitin. Formatting of the SafeAssign report is also very difficult to use effectively, reaching 104 pages, whereas the Turnitin report is appropriately formatted (40 pages)" (PSY)

Note for LTC Members: This does provide a useful basis on which to base the decision as to which TMS system University LTC should approve for investigative purposes – staff time is a precious resource and needs to be used efficiently.

Effectiveness

How well do the two TMS systems perform in terms of identifying 'matches'?

It is quite evident that the number of matches produced by SafeAssign and Turnitin vary greatly – with the latter producing a much higher 'hit rate' than the former. The differences in performance in identifying matches is actually quite 'stark' – as can be seen in the case of HUM below, where the Turnitin scores are often much higher. The POs have presented in their reports compelling and convincing evidence that use of Turnitin (as opposed to SafeAssign) would have had a **concrete and material impact** on the outcome of Plagiarism Meetings – i.e. the level of the offence, and the subsequent penalty imposed. This suggests that the use of Turnitin will add considerable value to the process of detecting and investigating cases of P & C in future and provides a sound basis on which to justify the subscription fees that the use of Turnitin will require.

Comments from POs:

"SafeAssign failed to detect blatant plagiarism in 7 of the 8 cases I reviewed. Only for 1 case did SafeAssign come up with anything like the same amount of matches" (ENV).

"Overall, Turnitin is more reliable and sensitive at detecting plagiarised content, frequently detecting sentences from journal articles that were omitted by SafeAssign, as well as from other sources (e.g. student papers). Turnitin also provides more information on the other sources, which is useful when checking and interpreting the highlighted content" (PSY).

"I put the Dissertation through SafeAssign which found a 19% match but these were mainly odd words & references but NOT the manually found identically copied text. The marker then spent a considerable length of time looking for more plagiarism but did not look manually through the whole document as it was a substantial piece of work and enough evidence had been discovered to undertake a Plagiarism meeting with the student…. I put the work through Turnitin and the work was found to have a 66% match with the markers identified areas found (attached). This… would have changed the outcome of the meeting from a Low offence to a Medium offence with the amount of Plagiarised material within the work" (HSC).

"I put the work through SafeAssign (24% match) which does identify some of the plagiarised work (attached). On putting it through Turnitin (57% match) which highlighted more comprehensively plagiarised work" (HSC).

"In terms of identifying matches, SafeAssign's performance shows a stark and in fact worrying disparity to the performance of Turnitin… With Turnitin it would have been extremely easy and effective to establish that the student had effectively resubmitted the plagiarised piece of work as reassessment. Because of SafeAssign, we had to go through the assignment manually page by page" (LAW).

"This is where the Turnitin software far exceeds the SafeAssign programme. SafeAssign picks up very little useful material from a social work perspective. In the five assignments I have put through both systems there is only one that is a close match. All the rest of the assignments had more relevant material highlighted using Turnitin" (SWK).

"Turnitin picks up more matches that SafeAssign. What takes time is recalculating / reducing the raw percentage in view of the unnecessary matches which must be ignored and which both pick up. SA makes more such matches and I find I alter the raw percentage for SA by 57% on average" (INTO).

"On running the work through SafeAssign it identified 15% plagiarism. These elements were all detected as from 'other students' papers.' On running the work through Turnitin 50% plagiarism was detected. The outcome for the student would have been materially different in reference to the 'extent of plagiarism' element of the classification guide (Curtin Grid)" (HUM).

"Turnitin performed better on some problematic essays as it identified matches with various internet sources that SafeAssign didn't pick up. I checked manually doing a Google search and the matches highlighted by Turnitin appeared to be correct" (NBS).

"In some instances SafeAssign matched text from the bibliography as well as names such as 'Shakespeare, William' that misrepresent the extent of plagiarism. This does not happen as Turnitin can filter the bibliography as well as small matches" (HUM).

"The matching performance of SafeAssign is extremely poor. This is not surprising as the database SA uses has overall a very limited number of legal journals; and the majority of the legal journals are not the ones generally read, used and referred to in UK HE. Turnitin, however, even accesses

the ebook repository of the OUP. The SafeAssign detection rate of sources freely available on the internet is also extremely poor" (LAW)

"Again, without the diligent work of the marker, this case would either not have been pursued or the level of plagiarism would have been misjudged. It needs to be noted that this student has been permanently expelled from UEA for this offence at SSDC level" (LAW).

"SafeAssign failed to pick up a number of matches from various websites. These are often sentences cut together from numerous sources – presumably to make the students' English appear better. Turnitin did identify these" (HUM).

"My personal experience has shown SafeAssign to be vastly inferior to Turnitin in terms of effectiveness in every single case I have investigated… I have been unable to utilise these SafeAssign reports in any plagiarism meetings as they do not pick up most of the plagiarised text, are not clear to interpret or explain to the student, and are not downloadable is a clear or usable format. As such they are unreliable to use as evidence and confusing for students. In each case I have generated a report, but have gone on to spend many hours manually searching for the evidence that SafeAssign has failed to discover, and to present it in a format that is clear to the student. In most cases the original source was easily available using Google!" (NBS)

"There is no doubt that reliance on SafeAssign rather than Turnitin would lead to a huge underestimation in the amount of plagiarism in individual cases and impact on the penalty imposed" (NBS).

"Turnitin seemed to be much better at detecting matches" (MED).

"Like chalk and cheese…. SafeAsssign regularly generates spurious matches of phrases such as "1.1 Introduction" and references in the bibliography. SafeAssign percentages and highlighting are either indecipherable or downright unhelpful and as a result I have avoided using it as evidence" (DEV).

"…the effectiveness of Turnitin in terms of detecting matches exceeds that of SafeAssign in every case I have examined" (MED).

"In one case, a marker identified plagiarism and because of the particular circumstances I was asked to check some previous work that the student had submitted as well as the one in question. Turnitin enabled me to identify that this student had committed serial plagiarism. It turned out the meeting that the student was having a number of problems that he needed help with. However, SafeAssign detected … almost none in the previously submitted ones. So without Turnitin I would not have found out that this student was a serial plagiariser" (ENV).

Note for LTC Members: The difference in performance between the two TMS systems is stark in terms of their effectiveness in identifying matches and showing the original sources. The difference in performance is sufficiently significant to ensure that this should be given considerable emphasis as a criteria for LTC to decide which TMS system to approve.

Matches
How effective is the software in detecting matches with text located in a) websites b) journal articles   c) Previously submitted student scripts d) Other online resources/databases?

Figures in Red (below) are those where the % score difference between SafeAssign and Turnitin is 30% or greater – i.e. examples where the 'gap' in performance and 'scores' are likely to have a significant impact on the level of the offence and the penalty imposed at a Plagiarism Hearing.

From HSC:
Case 1: SafeAssign; 3%,   Turnitin 11%   Difference: 8%   Range: 8-33%
Case 2: SafeAssign 19%,   Turnitin 66%   Difference: 47%
Case 3: SafeAssign 24%,   Turnitin 57%   Difference: 33%

From HUM:
Case 1: SafeAssign 15;   Turnitin 50%   Difference: 35%   Range: 26-36%
Case 2 SafeAssign 2.15%;   Turnitin 39%   Difference: 36.8%
Case 3 SafeAssign 3.9%;   Turnitin 30%   Difference: 26.1%

From LAW:

Case 1: SafeAssign 5%   Turnitin 62%   Difference: 57%   Range: 14-66%
Case 2 SafeAssign 9%   Turnitin 47%   Difference: 38%
Case 3 SafeAssign 16%   Turnitin 30%   Difference: 14%
Case 4 SafeAssign 7%   Turnitin 73%   Difference: 66%
Case 5 SafeAssign 56%   Turnitin 97%   Difference: 41%

From SWK:
Case 1: SafeAssign 4%   Turnitin 20%   Difference: 16%   Range: 4-35%
Case 2 SafeAssign 19%   Turnitin 31%   Difference: 12%
Case 3 SafeAssign 69%   Turnitin 73%   Difference: 4%
Case 4 SafeAssign 14%   Turnitin 49%   Difference: 35%
Case 5 SafeAssign 16%   Turnitin 46%   Difference: 30%

From MED:
Case 1: SafeAssign 16%   Turnitin 30%   Difference: 14%   Range: 14-46%
Case 2 SafeAssign 9%   Turnitin 47%   Difference: 38%
Case 3 SafeAssign 12%   Turnitin 58%   Difference: 46%
Case 4 SafeAssign   8%   Turnitin 51%   Difference: 43%
Case 5 SafeAssign   7%   Turnitin 27%   Difference: 20%

From LDC:

	SafeAssign	Turnitin	Difference	Range: -8% - 43%
Case 1	37	29	- 8%	
Case 2	27	70	43%	
Case 3	3	16	13%	
Case 4	1	27	26%	
Case 5	12	32	20%	

From NBS:

	Level	Turnitin	SafeAssign	Difference	Range: 3-49%
1.	UG	16%	13%	3%	
2.	PG	20%	17%	3%	
3.	PG	58%	22%	36%	
4.	PG	18%	3%	15%	
5.	UG	40%	9%	31%	
6.	UG	31%	4%	27%	
7.	PG	33%	1%	32%	
8.	UG	71%	22%	49%	
9.	UG	47%	13%	34%	
10.	UG	46%	3%	43%	

From HSC:
Case 1 – Turnitin reported 38%, SafeAssign reported 18%   Difference: 20%
Case 2 – Turnitin reported 44%, SafeAssign reported 11%   Difference: 33%

From ENV (TYN):

Assignment	SafeAssign	Turnitin	Difference	Range: 9-57%
1	7%	56%	49%	
2	1%	47%	46%	
3	5%	62%	57%	
4	11%	27%	16%	
5	15%	35%	20%	
6	7%	38%	31%	

7	4%	13%	9%
8	60%	68%	8%

Comments from POs:

"SafeAssign did not pick up the main source. Without the diligent work and will of the marker to go beyond the call of duty regarding the "manual" investigation the student would not have been prosecuted and found guilty of a high level case of plagiarism" (LAW).

"The main source that has been identified by Turnitin is a freely available PhD thesis on the internet. SafeAssign did not pick it up" (LAW).

"SafeAssign failed to pick up a significance match, where the submission appears to be a verbatim of a Vodafone mission statement. TI identified and highlighted this" (NBS).

"It is clear from the above figures that Turnitin routinely detects far more matched text than SafeAssign – relying on these scores would lead to very great differences in the penalty applied" (NBS).

"In total Turnitin found 22 matches from journals and other published articles, where as SafeAssign only found 3. Not only was Turnitin better at finding and identifying matched text, it presents it a straightforward and clear format which can easily be sent and/or explained to students" (NBS).

"Turnitin was much better at detecting text in websites and journal articles. The two systems were comparable in their effectiveness in detecting matches in previously submitted UEA student scripts" (MED).

"The evidence I have attached to this report speaks for itself. See particularly the essay from DEV-M070, in which Turnitin detected an 86% match, whereas SafeAssign could only manage 2%! In this case, Turnitin saved me literally hours of work" (DEV).

It is clear from the Reports provided by POs that the TMS systems perform very differently in terms of the 'matches' they identify, and in terms of the similarity 'Scores' that feature in the Originality Reports (ORs). The difference in scores is often more than 30 percentage points and differences greater than 40 percentage points are not uncommon. In the 16 case studies presented above, 10 exhibit a 'difference' in the OR 'score' of 30% points or more. In cases like these, the difference in performance is highly likely to have a direct and significant impact on the judgment made by the PO in a Plagiarism Hearing, and on the subsequent penalty imposed on the students in question.

Note for LTC Members: The difference in the effectiveness of the two TMS systems in identifying 'matches' with a wide range of original external sources is sufficiently pronounced to make a strong case for the adoption of Turnitin as the preferred TMS system at UEA for investigative purposes – and LTC would be well-advised to give this aspect of the evaluation considerable weight in its decision.

Suitability for batch screening

Although both of the TMS systems facilitate batch screening, and do so in much the same way, the greater performance of Turnitin as a 'screening tool' (i.e. one that identifies, quickly and easily, any issues for concern in students' scripts by providing an accessible, easily-understood, and concise insight into 'matches' and original sources etc) has led some POs to conclude that Turnitin would be the most effective tool for batch screening in the future. However, this would bring very significant implications in terms of building in the technical solutions that would ensure that its use was properly aligned with recent developments in online marking within the University.

Comments from POs:

"In view of dismal performance of SafeAssign in my tests, and also in the tests performed by CTEL in an ENV batch screening exercise, SafeAssign is completely unsuitable for batch screening" (ENV).

"My fear is that, given the discrepancy in the similarity index, the use of Safe Assign will give students a sense of false security. Also, it does not pick up Wikepedia entries, which, at least in LDC, are commonly used" (LDC).

"I would not feel confident in using SafeAssign for batch screening as this would give students false confidence about their writing abilities and/or their ability to use sources unacknowledged… Turnitin is a MUCH more reliable tool to use for Health students in identifying relevant sources that have been used/accessed/inappropriately referenced. It also identifies where students have poorly paraphrased… " (HSC).

"Because of the issues regarding SafeAssign's lack of user-friendliness I would strongly advise against the use of batch screening using it. It not only fails to assist POs and markers in their plagiarism investigation, but it has also caused confusion with students that were presented with Originality Reports generated by SA. The stated percentages and the way the alleged plagiarism is highly misleading and … has the potential to lead to an increase in student appeals" (LAW).

"I am reluctant about batch screening on the whole. But I would suggest that it would be almost pointless on SafeAssign as it would flag cases where students had referenced a lot. SafeAssign also fails to account for publications, something that Turnitin is much better suited for… SafeAssign does not accurately represent the extent of plagiarism in work submitted" (HUM).

"In batch screening the main issue would be to avoid Type II errors (allowing suspicious papers to pass under the radar). In this respect Turnitin did better because it flagged up more suspicious elements, some of which were relevant" (NBS).

"SafeAssign percentages cannot be relied on as an indication of the amount of matched text in a piece of work. It would be very dangerous and misleading for students to use this as a tool to indicate plagiarism, and impossible for markers to use the percentage scores generated as a guide" (NBS)

"I feel that Turnitin would be more likely to be effective for this purpose… I share the concerns expressed by my HSC colleague, that using SafeAssign might lead to the students having 'false confidence'" (HSC).

"SafeAssign's uninterpretable overall scores, and its abysmal detection capabilities make it unsuitable for 'batch' screening of coursework. We have shown here at least three cases (#2, #4 and #5) which would not have been detected by SafeAssign if used in screening mode, due to its failure to detect any significant matched text" (MED).

POs have presented compelling evidence in their reports that suggests that SafeAssign is poorly suited for use as a batch screening tool. This is reinforced by the outcome of the batch screening 'pilot' in MED, the results of which were reported to LTC at its meeting on 22 June 2016.

Note for LTC Members: The evidence provided by this comparative evaluation indicates that a small number of POs have a clear preference for the adoption of Turnitin as a system for batch screening. However, this needs to be subject to further discussion within the University since to use Turnitin in this way would involve considerable resource, cost and technical development implications.

Fitness for purpose
The primary purpose of TMS systems is to assist POs in determining whether students have infringed the University's Policy, and to determine how severe the infringement is (resulting in a particular penalty). It is important to know which of the two is most fitted for this purpose.

It is clear from the evidence provided by POs in their summary reports on the evaluation, that they have very serious concerns about the extent to which SafeAssign is fit for purpose. The evidence presented in previous sections of this report from the ADTP suggests that these concerns are well-founded. There does seem to be an unacceptably large difference between the performance of SafeAssign and that which would seem to be necessary to ensure a robust tool for investigating students' scripts for evidence of academic malpractice and infringements of the University's P & C Policy.

Comments from POs:

"SafeAssign is clearly unfit for purpose. Its outputs are also difficult to interpret…. On the other hand Turnitin is an excellent consistent tool for identifying plagiarism, useful in my work with markers where plagiarism is first identified, and also in preparing reports ahead of meetings, where the reports are essential in order for the POs to correctly assess the amount and nature of the plagiarism" (ENV).

"Overall I conclude that SafeAssign is not fit for purpose for investigation into, primary detection of, or education about, copied text in academic work. I believe that if the University considers this a sufficient tool for any of these purposes then its reputation and standards are at a very high risk… Any apparent cost-saving is false economy." (MED)

"When considering the fitness for purpose of the systems I have found it much easier and clearer to use Turnitin to identify where students' have infringed the University's P & C Policy. If I had been able to use Turnitin for Case 2 there would have been an increase in penalty imposed on the student" (HSC).

"In summary SafeAssign is not fit for purpose. It is time-consuming the use… it heavily dis-incentivises markers to report plagiarism; and impedes the work of Plagiarism Officers. POs become heavily reliant on the help of markers that have the expert knowledge in the academic field in question. The preparation of SSDC cases increases the required time even further. The manual highlighting and annotation of the plagiarised work became necessary in order to prove the plagiarism and present the case to the SSDC panel in an appropriate fashion that safeguarded the due process of the proceedings…. Relying on SafeAssign will result in a reduced detection rate for plagiarism; directly impacting on UEA's academic integrity. In the long-run this could have a damaging effect on UEA's international reputation and potentially impact on the recruitment of international students" (LAW).

"Turnitin would be more useful for me, though I depend on it less that do colleagues in the Schools" (INTO).

"In the cases discussed above it is clear that turnitin offers more reliable text matching. Both TMS require the PO to treat the work with care, reflection and nuance and a sensitivity to their own disciplines. That said the lack of academic publications consulted in SafeAssign, particularly in relation to the Humanities, makes it considerably less fit for purpose" (HUM).

"I do not believe that SafeAssign is fit for purpose. It is inconsistent, ineffective and does not generate a report that can be used as evidence in plagiarism meetings. Unlike SafeAssign, Turnitin is consistent, clear and user-friendly. It generates a report that is easy to interpret and can be downloaded in a format that can be provided as evidence for students" (NBS).

"There are also serious potential risks from legal challenges by students and damage to UEA's reputation. I would have to seriously question whether I could continue in my role as PO if the only text-matching tool at my disposal were SafeAssign" (NBS).

"Given my experience of using both systems, I much prefer Turnitin. It has a better success rate in identifying plagiarised work, and is clearer to read and therefore easier to explain with students" (HSC).

"I believe that SafeAssign is not fit for purpose in assisting POs in identifying matches from web pages and papers. Too many sources are missed. It also presents its findings in a less useful format" (MED).

"Turnitin is fit for purpose; SafeAssign is not. SafeAssign has not saved me any time or strengthened the case in any of the plagiarism cases I have investigated to date. It is, however, very effective at creating red herrings - matches in reference lists, random 3-4 word phrases, headings etc" (DEV).

"From my answers in 'Effectiveness' it is clear that I must conclude that Turnitin is overall much more effective and therefore more reliable" (LDC).

"Turnitin is undoubtedly more useful that SafeAssign in evaluating cases. SafeAssign has repeatedly failed to detect content that has been manually identified as plagiarised (i.e. by the marker/POs as part of the evaluation process), whereas Turnitin has successfully highlighted this content" (PSY).

Note for LTC Members: The evidence indicates that Turnitin would be by far the most effective of the two TMS tools for Investigative Purposes. In particular, the 'value-added' by the

Appendix 1
E mail sent to all Plagiarism Officers on 6 May 2016 confirming the process of the evaluation and the criteria against which the performance of the two TMS systems would be measured.

Dear Plagiarism Officers,

Further to my previous e mail of the 4 April, I would like to confirm at this stage how the parallel comparative evaluation of SafeAssign and TurnItIn will be carried out. All POs will be expected to participate in the evaluation.

The contact details of POs have been provided to me by LTS. If you are no longer Plagiarism Officer for your School, please forward this email to your successor, and alert Michele Pavey in LTS accordingly.

This calendar invitation sets out two possible dates (either of which you are free to attend) for an evaluation workshop when it will be possible to meet with Alicia McConnell, myself and members of Alicia's team in an IT lab to carry out a like-for-like comparison between Safe Assign and Turnitin, using real student scripts. Further workshop sessions may be scheduled if felt to be necessary.

The two dates of the two workshops are as follows:

Tuesday 17th May, 14.00 – 16.00 CSED IT Training Room in the LaRC (Library)

Friday 20th May, 13.00 – 15.00 CSED IT Training Room in the LaRC (Library)

[Note: a third workshop was added following the sending of the email]

Wednesday 25 May, 10:00-12:00 am *CSED IT Training Room in the LaRC (Library)*

These workshops will provide an opportunity to run-off Originality Reports using both systems, and to generate evidence that will inform a report by the ADTP to University LTC in July this year.

It is expected that LTC will make a final decision in July as to which of the two TMS systems will be adopted as the University's preferred system from 2016/17 onwards, both for investigative and batch screening purposes.

Taking-out a full year's subscription to Turnitin will cost £45,000. It is important, therefore, that the University can assure itself that this cost is sufficiently off-set by the additional 'value' that Turnitin can bring to the operation of our Plagiarism & Collusion Policy, and assisting in the detection of infringements of the Policy.

<u>The process:</u>

The evaluation process will proceed as follows:

1) Plagiarism Officers will have access, including passwords, to both TMS systems.

2) The evaluation will draw on student scripts which have already been the subject of a formal investigation by PO's (e.g. which formed the focus of a Plagiarism Hearing) and where the Originality Report was used a key from of evidence in determining guilt/innocence, the level of severity (and associated penalty).

3) POs will need to bring these sample scripts on a memory stick to the workshops highlighted above, or have access to them on a cloud drive. If POs do not currently have copies in their own possession, these can be secured via your LTS Hub Team Leader. Hub Team Leaders have been alerted in advance and will be able to support you in identifying appropriate scripts for inclusion in the evaluation.

4) It is up to PO's to select appropriate scripts. These could be UG or PGT level and may be from the current or previous academic years. Please ONLY draw on scripts which have already been subjected to Plagiarism Hearings.

5) POs are asked to bring a minimum of five and maximum of 20 scripts, in a digital format, to the workshop.

6) At the workshop PO's will be given guidance on how to submit the scripts to SafeAssign and to Turnitin, and how to access the Originality Reports.

7) POs will also be offered guidance on how to summarise the results of their like-for-like comparison using the proforma provided.

8) After the workshops POs will be able to reflect in more detail on the relative merits of the evidence contained in the respective originality Reports.

9) PO's will be asked to present their summary analysis using a standard proforma (see attached) including copies of, or links to the appropriate originality reports.

10) The PO's summary analysis, with Originality Reports, should be submitted to the ADPT no later than **10 am Wed 22 June 2016**. Send to: a.longcroft@uea.ac.uk

11) The ADTP will draw on the results of your collective analytical summaries in order to draft proposals for consideration by University LTC which set out the case for a preferred TMS solution for the foreseeable future.

Criteria

It is proposed that the evaluation focuses on the following criteria for establishing the relative merits/strengths of SafeAssign and Turnitin:

- *Speed* – How long does it take to run-off Originality Reports?
- *Ease of use* – How user-friendly is the software and how easy is it to interpret the Originality Reports generated?
- *Effectiveness* - How well do the systems perform in terms of identifying 'matches'?
- *Matches* – How effective is the software in detecting matches with text located in a) websites b) journal articles   c) Previously submitted student scripts d) Other online resources/databases?
- *Suitability for batch screening* – which system has functionality that is best suited for routine batch screening of student coursework submissions?
- *Fitness for purpose* – the primary purpose of TMS systems is to assist POs in determining whether students have infringed the University's Policy, and to determine how severe the infringement is (resulting in a particular penalty). It is important to know which of the two TMS systems provides evidence in a format which is aligned with this purpose – i.e. which is most helpful for POs in the execution of their responsibilities.

If POs are unable to attend either of the two workshops offered by the CTEL team, they are free to prepare a summary analysis in their own time and submit it to the ADTP by the same date (Wed 22 June 2016).
I look forward to receiving your completed reports.
Many thanks and best wishes.
Adam

Dr Adam Longcroft
Academic Director for Taught Programmes
Learning
University of East Anglia
Norwich
NR4 7TJ

Dr Adam Longcroft
Academic Director for Taught Programmes
Learning
University of East Anglia
Norwich
NR4 7TJ

Appendix 2
The report proforma (template) provided to POs by ADTP on 6 May 2016 for drafting their summary reports.

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016

The process:
The evaluation process will proceed as follows:
12) Plagiarism Officers will have access, including passwords, to both TMS systems.
13) The evaluation will draw on student scripts which have already been the subject of a formal investigation by PO's (e.g. which formed the focus of a Plagiarism Hearing) and where the Originality Report was used a key from of evidence in determining guilt-innocence and the level of severity (and associated penalty).
14) POs will need to bring these samples of students scripts on a memory stick to the workshops highlighted above, or have access to them on a cloud drive. If POs do not currently have copies in their own possession, these can be secured via your LTS Hub coordinator. Hub Coordinators have been alerted in advance and will be able to support you in identifying appropriate scripts for inclusion in the evaluation.
15) It is up to PO's to select appropriate scripts. These could be UG or PGT level and may be from the current or previous academic years. Please ONLY draw on scripts which have already been subjected to Plagiarism Hearings.
16) POs are asked to bring a minimum of five and maximum of 20 scripts, in a digital format, to the workshop.
17) At the workshop PO's will be given guidance on how to submit the scripts to SafeAssign and to Turnitin, and how to interpret the Originality Reports.
18) POs will also be offered guidance on how to summarise the results of their like-for-like comparison.
19) After the workshops POs will be able to reflect in more detail on the relative merits of the evidence contained in the respective originality Reports.
20) PO's will be asked to present their summary analysis using a standard proforma (see attached) including copies of, or links to the appropriate originality reports.
21) The PO's summary analysis, with Originality Reports, should be submitted to the ADPT no later than 10 am Wed 22 June 2016. Send to: a.longcroft@uea.ac.uk
22) The ADTP will draw on the results of your collective analytical summaries in order to draft proposals for consideration by University LTC which set out the case for a preferred TMS solution for the foreseeable future.

PLEASE NOTE: THE ADTP WILL BE ATTACHING COPIES OF THIS COMPLETED PROFORMA TO HIS REPORT TO UNIVERSITY LTC.
Please aim to keep the completed form to a word count limit of around 2,500 words.

Criteria
It is proposed that the evaluation focuses on the following criteria for establishing the relative merits/strengths of SafeAssign and Turnitin:
- *Speed* – How long does it take to run-off Originality Reports?
- *Ease of use* – How user-friendly is the software and how easy is it to interpret the Originality Reports generated?
- *Effectiveness* - How well do the systems perform in terms of identifying 'matches'?
- *Matches* – How effective is the software in detecting matches with text located in a) websites b) journal articles c) Previously submitted student scripts d) Other online resources/databases?
- *Suitability for batch screening* – which system has functionality that is best suited for routine batch screening of student coursework submissions?
- *Fitness for purpose* – the primary purpose of TMS systems is to assist POs in determining whether students have infringed the University's Policy, and to determine how severe the

infringement is (resulting in a particular penalty). It is important to know which of the two TMS systems provides evidence in a format which is aligned with this purpose – i.e. which is most helpful for POs in the execution of their responsibilities.

CRITERIA
Speed
Ease of use
Effectiveness
Matches
Suitability for batch screening
Fitness for purpose
Once completed, please forward this proforma to the ADPT no later than <u>10 am Wed 22 June 2016</u>. Send to: a.longcroft@uea.ac.uk **THANK YOU**

Appendix 3
Summary Reports from School Plagiarism Officers

Report from HSC (Anna Harris)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
From my experience the Speed of creating originality reports depends on the size of the document being uploaded and the time of day the document is uploaded to either SafeAssign or Turnitin. My recollection is that Turnitin can take slightly longer in the afternoon to upload but I have never had a problem or frustration with either system in this respect.
Ease of use
I undertook training in both SafeAssign and Turnitin but unequivocally find Turnitin the most easy to use. Uploading documents to each system is the same. Managing the information, interpreting the information, the visual presentation of the information and the overall satisfaction with Turnitin's performance is far and above that of SafeAssign in ease of use.
Effectiveness
This is where the greatest discrepancies occur. During last Term I had 2 cases to investigate and this Term 1 case.
Case 1 – this case was brought to me by the marker with suspicions from the 2 very different styles of writing within the work. The first part of the work was articulate and well sourced for a first year undergraduate. The latter part of the work was poorly written and the contrast was glaring (the work did not reflect the assessment criteria).
On putting the essay through SafeAssign it identified a 3% match which were key words and references (attached).
The marker's original email had flagged up 2 long reports (email attached) which I then read (as they did not come up in SafeAssign) in line with the document to identify and similarities. None were found. This was hugely time consuming as one paper was 381 pages.
At a later date on putting the work through Turnitin in my role as External Examiner the work raised an 11% similarity and the findings were acceptable (attached).
The Plagiarism meeting was very successful in that the student could explain the differences, demonstrate his work and difficulties and as a consequence measures were put in place to provide additional support for the findings.
Case 2 – this case was brought to me by the marker with suspicions following her identifying anomalies in the writing of a 3rd year Dissertation eg when the student referred to "…our nursing and midwifery degree…" referenced King's College London and was found to be from their website. The marker then looked more closely and found other direct sources of information not correctly referenced or attributed to an author.
I put the Dissertation through SafeAssign which found a 19% match but these were mainly odd words & references but NOT the manually found identically copied text (attached).
The marker then spent a considerable length of time looking for more plagiarism but did not look manually through the whole document as it was a substantial piece of work and enough evidence had been discovered to undertake a Plagiarism meeting with the student.
At a later date, in my role as an External Examiner I put the work through Turnitin and the work was found to have a 66% match with the markers identified areas found (attached). This was not used during the Plagiarism meeting but the outcome of the meeting would have changed the outcome of the meeting from a Low offence to a Medium offence with the amount of Plagiarised material within the work (>20%) (attached).
Case 3 – this case has just come to me for review and I am reporting on this prior to a formal Plagiarism meeting being held. The marker came to me with this students Dissertation as she had concerns about the quality and the marker's knowledge of the students' academic ability from being her PA. The marker highlighted some areas of the work and produced sources where the work had been taken from. I agree this requires further investigation.

I put the work through SafeAssign (24% match) which does identify some of the plagiarised work (attached). On putting it through Turnitin (57% match) which highlighted more comprehensively plagiarised work (attached).

Matches

Case 1:
SafeAssign match 3%, Turnitin match 11%
Case 2:
SafeAssign match 19%, Turnitin match 66%
Case 3:
SafeAssign match 24%, Turnitin match 57%
As stated above it was not just the % of the matches that were different but what the packages matched ie SafeAssign odd words & references, some relevant text, Turnitin actual passages direct from sources.

Suitability for batch screening

I would not feel confident in using SafeAssign for batch screening as this would give students false confidence about their writing abilities and/or their ability to use sources unacknowledged. Additionally the sources SafeAssign appear to access do not reflect the materials the students in HSC access eg Health related sources.
I appreciate Turnitin is not without its' nuances but I find it a MUCH more reliable tool to use for Health students in identifying relevant sources that have been used/accessed/inappropriately referenced. It also identifies where students have poorly paraphrased in that it picks out the majority of the work used omitting the students influences. This can then be cross referenced easily using the identification part of the Turnitin package where you can explore the source of the highlighted part.

Fitness for purpose

When considering the fitness for purpose of the systems I have found it much easier and clearer to use Turnitin to identify where students' have infringed the Universities Policy. As one case identifies above if I had been able to use Turnitin for Case 2 there would have been an increase in penalty imposed on the student.
Where a marker and PO are in agreement plagiarism has taken place the use of SafeAssign is much more time consuming and less accurate for HSC students assessments. I would have concerns about markers coming forward with their concerns about possible plagiarism if they become aware they will need to undertake a substantial amount of work manually researching the students' work to prove a case.

Report from LAW (Sven Gallasch)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
The time it takes to generate originality reports (OR) using SA or Turnitin is similar. Depending on time of day and the length of the submitted piece of work it might take between 1min and 1hr. Having said that, I believe that "speed" should not be regarded as a key criteria to establish the suitability and effectively the choice of the TMS used at UEA. The time that it takes to generate reports becomes irrelevant, if the generated report is of such poor quality that makes it unusable for further and more detailed investigation; which is the case for SA.
Ease of use
SA is not user-friendly at all, especially once the originality report is shared with the marker (for further investigation) or with the student (for the plagiarism meeting).
Effectiveness
In terms of identifying matches, SA's performance shows a stark and in fact worrying disparity to the performance of Turnitin.

SA is not user-friendly at all, especially once the originality report is shared with the marker (for further investigation) or with the student (for the plagiarism meeting).

The ease of use is directly linked to detected matches. The matches that SA does pick up (and these are hardly any) are not highlighted in full. The OR might highlight a single word and provides a percentage of the match once the highlighted area is clicked. Additionally, these percentages are highly misleading and often wrong. In a number of cases, detailed manual investigation has revealed that the picked up match is in fact is part of verbatim quote of the entire paragraph. In a number of cases the previous or following paragraph were also verbatim quotes from the same source, without having been identified as such by SA.

Turnitin, however, does highlight the entire paragraph in colour which is off tremendous help in identifying the extent of the potential plagiarism at a glance.

The situation becomes worse once the OR is shared with the marker or the student. The only way to share an OR generated by SA is by means of printing it as a PDF.

- In contrast to Turnitin, SA does not allow the print to be in colour.
- The report also does not keep the original formatting of the original submission which makes it almost impossible to compare the two documents. For example, footnote have the same font size as the main text and page numbers are extremely difficult to identify.
- In addition, all supposed matches are collated at the end of the report in little boxes that show excerpts from the original source. There is no way one can easily locate the alleged matches at the end of the document.

All these factors render the OR unusable for further investigation by the marker. The only person that has access to the remotely usable version in colour is the plagiarism officer. However, the PO is often not in the position to determine whether correctness of the main text and especially the appropriateness of the assigned footnotes, as the PO often lacks the required subject knowledge. As a consequence, we have stopped providing the students with the OR generated by SA, as the student will not only be unable to interpret the OR in any way; even more so the OR is often grossly misleading, as the match percentage at the top of the document is incorrect or at least unexplainable.

Instead we had to revert to manually highlighting the plagiarised parts of the assessment in the PDF file, also annotating the matching source. This was the only way we felt is available to ensure a due process for the student, especially in light of potential SSDC hearings and appeals to the SSDAC and OIA.

Effectiveness

In terms of identifying matches, SA's performance shows a stark and in fact worrying disparity to the performance of Turnitin.

Example 1: SA (5%) Turnitin (62%)

SA did not pick up the main source. Without the diligent work and will of the marker to go beyond the call of duty regarding the "manual" investigation the student would not have been prosecuted and found guilty of a high level case of plagiarism.

Example 2: SA (9%) Turnitin (47%)

The main source that has been identified by Turnitin is a freely available PhD thesis on the internet. SA did not pick it up.

Example 3: SA (16%) Turnitin (30%)

Although the percentage difference in this case is not as stark as in others, Turnitin picked up on the student's use of footnotes. In fact, every single footnote in this piece of assessment is directly lifted verbatim from a key source. Turnitin picked up the source. However, it needs to be said, that at the time of the investigation, the PO only had the OR generated by SA. Again, without the diligent work of the marker, this case would either not have been pursued or the level of plagiarism would have been misjudged. It needs to be noted that this student has been permanently expelled from UEA for this offence at SSDC level.

Example 4: SA (7%) Turnitin (73%)

SA did not pick up the two key sources used by the student in this assessment. Both are LLM dissertations that are freely available on the internet and can be found through google search. The marker identified only one of the sources through google search. The second source (which amounts to 24% of the student's work) would not have been identified without Turnitin.

Example 5: SA (56%) Turnitin (97%)

With Turnitin it would have been extremely easy and effective to establish that the student had effectively resubmitted the plagiarised piece of work as reassessment. Because of SA, we had to go through the assignment manually page by page.

Matches

The matching performance is extremely poor (at least in Law). This is not surprising as the database SA uses has overall a very limited number of legal journals; and the majority of the legal journals are not the ones generally read, used and referred to in UK HE. Turnitin, however, even accesses the ebook repository of at least OUP.

I understand from an email sent to all POs by Alicia Macconnell that SA's focus is supposed to be on freely available internet sources and our own internal year-on-year database of students' work. I cannot comment on SA's performance of detecting collusion as I did not have to deal with such cases this year. In contrast to Turnitin, SA does not pick up similarities to assignments submitted at other UKHE institutions. Unfortunately, this is not surprising as the University of West of England is the only UKHE institution that uses SA exclusively.

However, the detection rate of sources freely available on the internet is also extremely poor. The main sources used in Example 1, 4, 5 (above) were freely available and in fact were identified by the marker by means of a cursory google search, typing in single sentences or short paragraphs.

Suitability for batch screening

Based on my experience and especially because of the issues regarding SA's lack of user-friendliness (outlined above) I would strongly advise against the use of batch screening using SA. It not only fails to assist POs and markers in their plagiarism investigation, but it has also caused confusing with students that were presented with ORs generated by SA. The stated percentages and the way the alleged plagiarism is highly misleading and in my opinion has the potential to lead to an increase in student appeals.

Fitness for purpose

In summary in can only be concluded that SA is not fit for purpose. The remainder of this section outlines the potential detrimental effect the use of SA as TMS could have on the detection and investigation of plagiarism at UEA.

1. **Using SA heavily dis-incentivises markers to report plagiarism**:
   Due to poor quality of the OR generated by SA, I often had to ask markers to go manually through the suspected pieces of assessment, highlighting and annotating the plagiarised sections. This is extremely time-consuming and burdensome especially in light of the 20 day turnaround rule for feedback.
   This situation will be exacerbated following the proposed changes to the green book and the aspiration to a 15 day turnaround. If the adherence to this tighter turnaround becomes part of the promotion criteria (as suggested), members of staff have no incentive to flag plagiarism, as keeping to the 15 days becomes the paramount objective over the detection of plagiarism.
   For example, one of our staff members used **two days during his annual leave** (overseas holiday in the US) to manually highlight a plagiarised piece of assessment for an upcoming SSDC hearing. Although such a commitment to UEA's academic integrity is commendable,

it is at the same time an unacceptable burden and goes way beyond the call of duty. [Example 1 above].

2. **Using SA heavily impedes the work of POs:**
The time required for the investigation of plagiarism increased dramatically because of the use of SA. POs become heavily reliant on the help of markers that have the expert knowledge in the academic field in question. The preparation of SSDC cases increased the required time even further. The manual highlighting and annotation of the plagiarised work became necessary in order to proof the plagiarism and present the case to the SSDC panel in an appropriate fashion that safeguarded the due process of the proceedings; thereby reducing the student's chances to successfully appeal the outcome at SSDAC or OIA level. In the light of this, it also needs to be kept in mind that most of the POs are on ATR contracts. Fulfilling the role as PO to an appropriate standard significantly encroaches on the individual's time for research; the key performance criteria ATR members are judged against for career progress.

3. **Ultimate consequence is diminished detection rate:**
The combination of the two above mentioned factors is likely to lead to a reduced detection rate for plagiarism directly impacting on UEA's academic integrity. In the long-run this could have a damaging effect on UEA's international reputation and potentially impact on the recruitment of international students.

Report from HUM (Victoria Cann)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Ease of use
I find both systems easy to use in terms of upload. Turnitin, however, allows for greater ease in presenting the results of detected plagiarism.
Effectiveness
The following cases from IIH have been run through TMS and have received the following results: **Case 1** – this assessment was marked by me, I noticed different styles of writing and some quite sophisticated overviews of feminism for work of this level (FY). On running the work through SafeAssign (with deputy PO) it identified 15% plagiarism (satori1). These elements were <u>all</u> detected as from 'other students' papers.' On running the work through Turnitin (with deputy PO) 50% plagiarism was detected (CANN1). The results showed the presence of a range of online sources from which work was cut and paste. **The outcome for the student would have been <u>materially different</u> in reference to the 'extent of plagiarism' element of the classification guide.** **Case 2** – this work was brought to me by the marker who has suspicions relating to the consistency of writing and referencing with switches between in-text and footnoting as well as omissions from the bibliography. On running this work through SafeAssign a result of 15% was returned (tom4). All of these results were sourced from other students' papers. On running this work through Turnitin 39% of the work was detected to plagiarism (TOM4). This included internet sources and publications. **The outcome for the student would have been <u>materially different</u> in reference to the 'extent of plagiarism' element of the classification guide.** **Case 3** – this work was brought to me by the marker who has suspicions relating to inconsistencies in the writing style. On running this work through SafeAssign a result of 9% was detected (satom2). Sources included mostly other students' papers as well as detecting plagiarism in the bibliography. On running the work through Turnitin a case of 30% was returned (TOM3). These sources included internet sources, publications and student papers. **The outcome for the student would have been <u>materially different</u> in reference to the 'extent of plagiarism' element of the classification guide.** In all cases Turnitin more clearly communicated the cases of plagiarism as well as detecting instances of plagiarism.
Matches
Case1: 15% (SafeAssign); 50% (Turnitin) **Case 2:** 15% (SafeAssign); 39% (Turnitin) **Case 3:** 9% (SafeAssign); 30% (Turnitin) In some instances SafeAssign matched text from the bibliography as well as names such as 'Shakespeare, William' that misrepresent the extent of plagiarism. This does not happen as Turnitin can filter the bibliography as well as small matches.
Suitability for batch screening
I am somewhat reluctant for batch screening on the whole. But I would suggest that it would be almost pointless on SafeAssign as it would flag cases where students had referenced a lot. SafeAssign also fails to account for publications, something that Turnitin is much better suited for. In the cases discussed above it suggests that SafeAssign does not accurately represent the extent of plagiarism in work submitted.
Fitness for purpose
In the cases discussed above it is clear that turnitin offers more reliable text matching software. Both TMS' require the PO to treat the work with care, reflection and nuance and a sensitivity to their own disciplines. That said the lack of academic publications consulted in SafeAssign, particularly in relation to the Humanities, makes it considerably less fit for purpose.

Report from SWK (Christine Cocker)

Speed – How long does it take to run-off Originality Reports?
Both have a similar speed. I don't need the software to analyse large batches of information at once, so the speed of both to analyse scripts I upload is fine.

Ease of use – How user-friendly is the software and how easy is it to interpret the Originality Reports generated?
The Turnitin software is easier to interpret. It is colour coded and this translates onto the pages which are then able to be downloaded and copied for the student to examine as well. This is not the same for the SafeAssign system.

Effectiveness - How well do the systems perform in terms of identifying 'matches'?This is where the Turnitin software far exceeds the Safeassign programme. Safeassign picks up very little useful material from a social work perspective. In the five assignments I have put through both systems there is only one that is a close match (IT). All the rest of the assignments had more relevant material highlighted using 'Turnitin'. Please be aware that some of the material highlighted by both systems was not relevant

Five plagiarism cases in social work 2014-2016

Name	% SafeAssign matches	% Turnitin matches
Case 1	4%	20%
Case 2	19%	31%
Case 3	69%	73%
Case 4	14%	49%
Case 5	16%	46%

Matches – How effective is the software in detecting matches with text located in a) websites b) journal articles   c) Previously submitted student scripts d) Other online resources/databases?
Turnitin is better at picking up information from social work journal articles and books which have an on-line presence. Both picked up data from websites. SafeAssign does not appear to have any UK or USA based social work journals in its search range.

Suitability for batch screening – which system has functionality that is best suited for routine batch screening of student coursework submissions?
I do not use this function so I cannot comment on either programme's suitability for this task.

Fitness for purpose – the primary purpose of TMS systems is to assist POs in determining whether students have infringed the University's Policy, and to determine how severe the infringement is (resulting in a particular penalty). It is important to know which of the two TMS systems provides evidence in a format which is aligned with this purpose – i.e. which is most helpful for POs in the execution of their responsibilities.
It is my view that turnitin is the better programme for my purposes as a plagiarism officer in the School of Social Work. The cases that I have dealt with during the last academic year when only Safeassign was available, relied solely on the skill of the marker in identifying text from essays which closely matched material that they were aware of from their own reading/resources. SafeAssign did not pick up any relevant material, including material that the marker originally identified. This makes it virtually useless as a detection aid, and I could not rely on it at all to support my work.

Report from NBS (Peter Ormosi)

All scripts used for these tests had been run through SA before. I looked at three factors: speed, accessibility and effectiveness.

Test 1
- Speed: Didn't notice a significant difference in the upload time. SA is slightly easier to upload (as it doesn't require logging the name of the student and the name of the paper)
- Output accessibility: I personally don't see much difference in the accessibility of either output.
- Effectiveness:
    - SA failed to pick up a significance match, where the submission appears to be a verbatim of a Vodafone mission statement. TI identified and highlighted this.
    - TI identified more matches with essays submitted in other universities – however none of these were relevant.

I think altogether on this particular script TI did better.

Test 2
- Speed: Didn't notice a significant difference in the upload time.
- Output accessibility: I personally don't see much difference in the accessibility of either output.
- Effectiveness:
    - SA failed to pick up a number of matches from various websites. These are often sentences cut together from numerous sources – presumably to make the students' English appear better. TI identified these.
    - Again, a large number of other student essays identified – none of these were relevant.

I think altogether on this particular script TI did better.

Test 3

This was a low-match case under both SA and TI. The two applications performed equally well.

Test 4

Both SA and TI made a Type I error here (identifying a script as suspicious even though it isn't). There was no difference in the performance of the applications.

Test 5

SA and TI performed comparably. There were some false alarms of copying from other student essays but the possibility of this could be excluded.

Summary

Effectiveness: TI performed better on some problematic essays (see Test 1 and 2) as it identified matches with various internet sources that SA didn't pick up. I checked manually doing a Google search and the matches highlighted by TI appeared to be correct.

User friendliness: I did not find any difference in accessibility, speed or how user-friendly the applications were. I had no prior experience with either. I would safely say that based on this I would be happy to use either application.

Batch screening ability: In batch screening the main issue would be to avoid Type II errors (to allow suspicious papers pass under the radar). In this respect TI did better because it flagged up more suspicious elements, some of which were relevant.

Total: TI won based on being more effective on these five tests.

Report from INTO (Bill Horncastle)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
Speed
I've found no important difference in speed between the two.
Ease of use
My own competence in saving / sending / printing these reports is not enough for me to judge.
Effectiveness
We can't expect perfection. Tnt picks up more matches that Sfn. What takes time is recalculating / reducing the raw percentage in view of the unnecessary matches which must be ignored and which both pick up. Sfn makes more such matches (see attached table) and I find I alter the raw percentage for Sfn by 57% on average. Such adjustments for Tnt amount to 18%.
Matches
I counted the frequency of matches from articles, the web, and students' papers, though not the number within each category. Tnt and Sfn both found matches of all three types, though there may be some significance in Tnt sourcing articles for 5 students (out of 11) compared to 2 for Sfn.
Suitability for batch screening
I've never done this. An INTO colleague finds it easy with Sfn. UEA colleagues tell me only MED has done it with Tnt.
Fitness for purpose
Tnt would be more useful for me, though I depend on it less that do colleagues in the Schools. The majority of INTO students resort to having their work re-written or translated, and the software can't find that. (Not yet; voice recognition software distinguishes between one person and another, and writing also carries individual markers. Isn't this a research / enterprise area for UEA?

Report from NBS (Sarah Allen)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
Speed
At quiet times, both perform at a similar speed. Occasionally there is some delay while Turnitin generates its report, but this has never been a problem for me as I usually submit a script and then return to the completed report later. I have never been in a position where I need a text-matching report immediately.
Ease of use
It is marginally more straightforward to generate a SafeAssign report, but this is mainly due to the fact that it has fewer useful options: the ability to filter out very short phrases, and exclude bibliographies and sections in quotation marks etc In all other respects Turnitin is easier to use. It is very simple to understand and interpret the report, and therefore to explain to students.   The Turnitin report highlights the exact words which have been matched and links these to the source where they were found. The different matched sources are colour-coded, numbered, and listed in the order of largest percentage of matched text. SafeAssign does none of these things! Instead it highlights whole sentences which contain some matched text – it is not clear from the report which words have been matched. While the online SafeAssign report is in colour, it is not possible to download a colour version to show or even email to the student. The list of sources at the end is not numbered so can be very confusing where there are many different sources, and almost impossible to explain to a student in a meeting. The percentage score generated by Turnitin is crude, but simple and consistent.   It simply highlights the matched text and generates a percentage score which reflects the proportion of the whole text that has been matched. This is clear and easy to explain to students, and also to compare between different scripts. It is possible to filter out short phrases, reference lists, and quotations. This option is not available in SafeAssign. The percentages generated in SafeAssign are far less clear. The overall percentage score is presumably an indication of the proportion of text matched, as with Turnitin. However, the way in which SafeAssign highlights text (whole sentences highlighted where only part of the sentence actually matches; its inability to highlight all the text from a single source; no way of excluding correctly quoted text, small phrases or bibliographies) render this percentage score meaningless. Each section of matched text (listed at the end of each report) also generates a percentage score. The rationale for this is even more obscure. It appears from the instructions on Blackboard that this score is SafeAssign's estimate of the likelihood of that particular phrase having been plagiarised. How it makes this estimate is unclear, and how I could explain this to a student is even less clear!
Effectiveness
My personal experience has shown SafeAssign to be vastly inferior to Turnitin in terms of effectiveness in every single case I have investigated. During the past year I have investigated a large number of cases of suspected plagiarism as NBS is the largest school at UEA, and for most of that year I have only had access to SafeAssign as a software matching tool. I have been unable to utilise these SafeAssign reports in any plagiarism meetings as they do not pick up most of the plagiarised text, are not clear to interpret or explain to the student, and are not downloadable is a clear or usable format. As such they are unreliable to use as evidence and confusing for students. In each case I have generated a report, but have gone on to spend many hours manually searching for the evidence that SafeAssign has failed to discover, and to present it in a format that is clear to the student. In most cases the original source was easily available using Google! There is no doubt that reliance on SafeAssign rather than Turnitin would lead to a huge underestimation in the amount of plagiarism in individual cases and impact on the penalty imposed. This has serious implications for plagiarism officers trying to do an effective job at UEA in terms of time spent investigating cases, consistency of outcome, and fairness to students. It could also lead to potential legal challenges by students.

Matches
As part of this study, I **randomly** selected 10 scripts that I have investigated over the past couple of years, and generated a second report for each, using whichever software had not been used in the original investigation. The percentage scores are as follows: Level   Turnitin   SafeAssign 1.   UG   16%   13% 2.   PG   20%   17% 3.   PG   58%   22% 4.   PG   18%   3% 5.   UG   40%   9% 6.   UG   31%   4% 7.   PG   33%   1% 8.   UG   71%   22% 9.   UG   47%   13% 10.   UG   46%   3% It is clear from the above figures that Turnitin routinely detects far more matched text than SafeAssign, with some very significant differences in all but the first two cases – relying on these scores would lead to very great differences in the penalty applied. SafeAssign (see examples provided) tends to match items in the reference list, commonly used phrases and subject headings. The vast majority of matched text in the case studies consisted of very short phrases and even single words. Where it did match text from a particular source, it highlighted only part of the copied text (commonly just a few phrases) and frequently missed other text from the same source. This inconsistency has serious implications for fairness in detection It also became apparent from the 10 case studies I used that Turnitin had a better range of sources in its database: In total Turnitin found 22 matches from journals and other published articles, whereas SafeAssign only found 3. Not only was Turnitin better at finding and identifying matched text, it presents it a straightforward and clear format which can easily be sent and/or explained to students.

Suitability for batch screening
I am unable to comment fully on this as I have no experience of using batch screening. My only comment would be that SafeAssign percentages cannot be relied on as an indication of the amount of matched text in a piece of work. It would therefore be very dangerous and misleading for students to use this as a tool to indicate plagiarism, and impossible for markers to use the percentage scores generated as a guide. They would need to carefully look at each piece of work individually, which is what they do now – so it rather defeats the purpose of batch screening in the first place!

Fitness for purpose
I do not believe that SafeAssign is fit for purpose. It is inconsistent, ineffective and does not generate a report that can be used as evidence in plagiarism meetings. Turnitin is not infallible, but its worst fault is that it is occasionally slow and sometimes fails to detect a source. Its failure to detect a source is far less frequent than SafeAssign, whose database clearly does not include the types of journal articles that we expect our students to be using at this level of study. Unlike SafeAssign, Turnitin is consistent, clear and user-friendly. It generates a report that is easy to interpret and can be downloaded in a format that can be provided as evidence for students. Importantly, it can also be used in a formative way to help students understand about referencing and plagiarism. It is important to stress that in many cases plagiarisms need to be conducted remotely by Skype or even phone, particularly where students have finished their course and returned home. In these situations it is essential, in the interests of natural justice, that the student has sight of the evidence against them in a clear and accessible format, especially where English is not their first language. I would not feel happy using a SafeAssign report in this situation.

Whilst I appreciate there is a significant cost associated with the reintroduction of Turnitin, I believe there are even greater hidden costs associated with the sole use of SafeAssign. These include the significant additional time taken by plagiarism officers investigating and evidencing cases, the likelihood of missing many cases and therefore failing to educate students about referencing and deter them from plagiarising.

There are also serious potential risks from legal challenges by students and damage to UEA's reputation. I would have to seriously question whether I could continue in my role as PO if the only text-matching tool at my disposal were SafeAssign.

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
I have found both tools reasonably fast to use. There can be a delay of a few minutes with Turnitin, whilst it generates the report, and this shows a message which states that they are experiencing a busy period. I find that if I press the refresh button, this is usually resolved straight away.
Ease of use
I have found both systems reasonably easy to use, once I became familiar with them. Turnitin reports are clearer and easier to read than the SafeAssign reports.
Effectiveness
I will use 2 cases to compare the effectiveness. These are my two most recent cases, both of which have come to me during the spring semester. Both were dissertations. Both were brought to me by the marker, who had concerns that there may be plagiarism.   In both cases, Turnitin was far more effective. It identified a larger percentage of matches, and although some may be disregarded for both Turnitin and SafeAssign, due to identification of commonly used phrases, Turnitin identified longer phrases/sentences which SafeAssign only parts of.   On occasion when I have used SafeAssign, a significant amount of additional material, not identified by SafeAssign, was identified by myself, via (time consuming) manual searching.
Matches
Case 1 – Turnitin reported 38%, SafeAssign reported 18%   Case 2 – Turnitin reported 44%, SafeAssign reported 11%
Suitability for batch screening
I feel that Turnitin would be more likely to be effective for this purpose. I do have some reservations about then practicalities of implementing batch screening with our large student cohorts. I do share the concerns expressed by my HSC colleague, that using SafeAssign might lead to the students having 'false confidence'.
Fitness for purpose
Given my experience of using both systems, I much prefer Turnitin. It has a better success rate in identifying plagiarised work, and is clearer to read and therefore easier to explain with students. It is also less time consuming. I have concerns regarding the lack of material that is identified when using SafeAssign.

Report from MED (Jane Skinner)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
For the small sample used for this exercise, the originality reports were run off quite quickly. From experience, Turnitin can be a bit slow at times of peak demand.
Ease of use
I found the two packages roughly the same in terms of ease-of-use in generating an originality report. The Turnitin reports, however, are much easier to interpret. The different sources are colour coded. Verbatim matches are highlighted, so you can see which words were changed. The Turnitin report matches are logically numbered in the order they appear in the text. In contrast, SafeAssign highlights contiguous blocks of text, and reports a percentage match of dubious utility for the whole block, so you can't see exactly what has been changed. The order in which the sections of matched text appear at the end of the report seems arbitrary and different from the order in which they appear in the text.
Effectiveness
Turnitin seemed to be much better at detecting matches. See next section.
Matches
Turnitin was much better at detecting text in websites and journal articles. The two systems were comparable in their effectiveness in detecting matches in previously submitted UEA student scripts.   For example, in a portfolio report Turnitin found 47% of the text to be matched, to a variety of papers from websites and papers, whereas SafeAssign identified only 4%. The matches in SafeAssign are all from students at other institutions, who have also used the same source texts – i.e. SafeAssign is not finding the original material on the web and in papers, although it is easily available. If this portfolio report had been used in a plagiarism case, the originality reports would have made the difference between a "low" (from SafeAssign) and probably a medium (from Turnitin).   In another case, Turnitin's more extensive database of work from other UK institutions is demonstrated by another portfolio report that was matched at 97% to a paper submitted to the University of Leeds, whereas the SafeAssign database didn't have this paper.   All MBBS year 1 portfolio reports were submitted to SafeAssign in their summative version this year (to check "for real" for possible plagiarism). A random sample of 15 of these were submitted in their formative version, as part of this evaluation exercise, to check for matching different versions of the same work. The same pattern of submissions (summative then formative) was repeated in Turnitin. The "matching" statistics for the two versions are shown below for the two packages:

No.	Turnitin	SafeAssign
1	45%	62%
2	88%	100%
3	90%	100%
4	43%	65%
5	50%	62%
6	43%	56%
7	37%	45%
8	50%	45%
9	55%	74%
10	96%	100%
11	58%	69%

12	98%	100%
13	35%	53%
14	54%	76%
15	76%	100%

As can be seen, the SafeAssign matching scores are consistently higher. All of the students in the sample had in fact changed their documents at least somewhat between the formative and the summative, so it is not clear why SafeAssign is showing matches of 100%.

Attached are originality reports for "Y1fPR15-6 015" As can be seen, the texts are not identical, though reported by SafeAssign as such.

Suitability for batch screening
The method for batch screening wasn't that obvious in either package, but fairly easy to do once it had been demonstrated.
Fitness for purpose
I believe that SafeAssign is not fit for purpose in assisting POs in identifying matches from web pages and papers. Too many sources are missed. It also presents its findings in a less useful format. SafeAssign could be considered fit for purpose in identifying matches from previously submitted UEA coursework, but it is something of a blunt instrument, with a tendency to over-report the matching.

Report from DEV (Gareth Edwards)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
Comparable. Turnitin seems to take slightly longer during periods of high demand, but this is vastly outweighed by the usefulness and customizability of the reports.
Ease of use
Turnitin is vastly more user-friendly, and the originality reports it generates are much easier to use. Text highlighting is more accurate and clearer, and the ability to generate a PDF originality report which maintains the original file format is much more useful than the SafeAssign approach which renders the text as a HTML file, making it double-spaced in the process and losing formatting. Because Turnitin keeps the formatting the same, it is easier to compare the originally submitted version with the originality report.
Effectiveness
Like chalk and cheese. Where SafeAssign is chalk and I am a cheese-lover. SafeAsssign regularly generates spurious matches of phrases such as "1.1 Introduction" and references in the bibliography. SafeAssign percentages and highlighting are either indecipherable or downright unhelpful and as a result I have avoided using it as evidence. The evidence I have attached to this report speaks for itself. See particularly the essay from DEV-M070, in which Turnitin detected an 86% match, whereas SafeAssign could only manage 2%! In this case, Turnitin saved me literally hours of work.
Matches
Turnitin is vastly superior to SafeAsssign, both in terms of matching text from appropriate sources, allowing POs to exclude matches from various sources, and presenting the text matched in a format which makes it easy for POs to interpret. Turnitin is also more customizable, with options for instance of excluding text located in a bibliography/reference list. The database searched by SafeAssign only contains one journal in my field with full text access (Progress in Human Geography). Better coverage can be obtained simply doing a Google search within the UEA IP range, or Web of Science/Knowledge.
Suitability for batch screening
I am unable to comment on this criterion as I have not evaluated either in a batch screening situation
Fitness for purpose
Turnitin is fit for purpose; SafeAssign is not. SafeAssign has not saved me any time or strengthened the case in any of the plagiarism cases I have investigated to date. It is, however, very effective at creating red herrings - matches in reference lists, random 3-4 word phrases, headings etc. Plagiarism Officers perform an important role in maintaining academic standards, and the university should make the financial investment required to support them in this process. Indeed, the opportunity cost of time lost due to the extra manual searching SafeAssign has necessitated is significant, as that time could have been spent on other facets of the academic role, notably research and teaching.

Report from MED (Gill Price)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016 CRITERIA
Speed
Speed of working of either software is not perceived to be an issue. If response from either software is slow I do something else and go back to it. The only situation when this could be an issue is during training sessions, and to date we have not found this to be a major problem. Both systems are US-based and so tend to be slower later in the day when US usage increases. The major extra time cost of SafeAssign above Turnitin is in interpreting the originality report, determining the exact words in a highlighted sentence which have been copied, and in doing further checks (eg using Google phrase by phrase) when it is clear that SafeAssign has not detected all of the copied text in the document.
Ease of use
Submitting documents to SafeAssign is marginally easier than to Turnitin, but that is because there are fewer options in SafeAssign, eg. for creating sub-folders, setting dates of an assignment, excluding checking of the bibliography or small matches. For the purpose of this evaluation I set Turnitin to include matches to bibliographic material, to equate with SafeAssign's unmodifiable setting. Accessing originality reports (ORs) online is easy in both softwares – though if one wanted to see ORs after screening a cohort and did not want to see students names the access is much more convoluted and laborious. Within the online OR, both systems allow one to click on a highlighted match in order to see an excerpt of the indicated source. In some cases one can go through to the original source (as distinct from an excerpt or a web-copy of the source) in both softwares. Highlighting of matches is in colour in both softwares when viewed online. Downloading an OR in SafAssign has to be done via a printing icon, and print to PDF – there is no other way of downloading an OR to store offline. I have not seen any instructions about this from BB. Downloading can take some time in either software but this depends on other factors (including length of the report and number of matches, and Web activity at the time, and there is no clear difference between them in this regard. In downloaded ORs the formatting of the original script and the colour highlighting is maintained in the Turnitin downloaded version, neither of which occurs in SafeAssign. Numbering of matches is done in SafeAssign to indicate the source on a sentence by sentence basis, but no numbering is used to identify the 'matched text' excerpts which show the extent of the matching for each sentence against the suggested source. Interpreting whether a highlighted sentence does in fact constitute a match which could be plagiarism requires separate study of all these individual excerpts, and reconstitution to a whole so that tracts of copied text can be identified, a mind-boggling and senselessly time-consuming task. Having no link between the highlighted sentence in the text and the juxtaposed suggested source at the end of the OR is a very careless omission on the part of the SafeAssign designers. This makes it very difficult to navigate the matches – both in investigation and when using in a meeting with a student. The usefulness for teaching purposes is also severely compromised by these difficulties in presentation and navigation in SafeAssign reports. Interpretation of highlighted text is very simple with Turnitin, as highlighting is done on a word-by-word basis.   If a chunk of text is copied from a source and then individual words edited, Turntin will show the chunk of text as a whole (highlighted and numbered as an integrated source), with those individual words left un-highlighted. I have verified this by studying the Turnitin report against several submitted scripts and their indicated sources. This is critically important when determining how a source has been used, whether an insufficient attempt at paraphrasing on the part of a novice-writer or a purposeful attempt to short-cut the academic

process. The example of the dissertation chapter 2 (Case #5) highlights this, in which words describing the topic of the dissertation were substituted for the (entirely different) topic of the original source in a paper describing the results of a systematic literature review. This is important in determining the **nature of a case** in the Plagiarism policy grid – indicating the extent to which the assessment has been jeopardised. Turnitin shows clearly where spelling has been changed in a copied chunk of text (eg from US to English spelling) – the changed words are not highlighted. This can also have a bearing on the judgment about **intention to deceive** in a case.

Another important aspect of the method of highlighting matches is the ability reliably to detect plagiarism of secondary sources. On the face of it a piece of text may appear to be well referenced, but submitting to Turnitin would often highlight not only chunks of that text but also the in-text citations. This indicates reliably that another source which reported from and referred appropriately to the original has been copied verbatim. When SafeAssign highlights the entire sentence (including the in-text citation) it appears as though such plagiarism has occurred, until one realises that this is merely the idiosyncrasy of the software. An example of this is on page 4 of the SafeAssign report on Case #4.

Identifying and tracing the sources of matched text can be difficult and time-consuming with both systems. Turnitin's most infuriating features during investigations are found here: websites containing copies of journal articles and other websites, and previously submitted student papers are often listed as sources. There are ways around this, but they add considerable time to an investigation, which can take days.
Turnitin has another format for its reports – an http version – which tends to list the sources more usefully than the same report in formatted 'document view' – eg. the title and authors of published articles are more likely to show here, whereas in 'document view' the sources tend to be listed as websites bearing a copy of at least part of the published text.

Interpretability of the scores produced in OR's are also an important aspect of ease of use. The 'similarity score' by Turnitin is presented as the proportion of the text which has been matched with another source. It seems to correspond with the volume of highlighted text. Neither the 'overall SafeAssign score' nor the 'sentence-matching score' produced by SafeAssign are interpretable – Blackboard gives conflicting information, none of which can yet be verified. On a downloaded OR it labels the information as a "X% match". However, this overall score does not always seem to correspond with the volume of highlighting, even when multiplied by the sentence overall scores for each sentence. Due to this lack of interpretability SafeAssign's usefulness for meetings with students, for determining the **extent of plagiarism**, and for operation of 'batch' screening of coursework, is severely compromised.

Effectiveness
The effectiveness of both systems is illustrated by the cases purposively sampled and described under 'Matches' below. I chose the cases to illustrate different aspects of submitted work, and the likely sources used in constructing them, which I have had to deal with over the past 7 years as Plagiarism Officer for the Medical School.   Neither system is perfect – case #1 illustrates one for which neither Turnitin nor SafeAssign detected a central source – but the effectiveness of Turnitin in terms of detecting matches exceeds that of SafeAssign in every case I have examined. This happens to be reflected in the scores presented by each software, but my judgement is based not on the scores but on the matching of text wtihin the reports. SafeAssign typically matches material in the bibliography of a submitted document, and fragments of sub-headings and in-text citations in the text, and inadequately and inconsistently detects important copied material in the script.

Matches			
Case number and type	**Turnitin– Similarity index**	**Safe-Assign Overall score**	**Comments**

1. MSc Project proposal (PGT) – see annotated Turnitin-report	30	16	The student copied extensively, with partial editing, from a set of slides in the 'Slideshare' website (found by the marker), as well as from other online sources. Neither Turntitin nor SafeAssign detected the copying from 'Slideshare'. For the case I had to annotate the Turnitin report (attached). SafeAssign highlighted part of the first web source used, and another web source (NICE), both found by Turnitin, but did not detect two other paragraphs found by Turnitin.
2. MBBS UG Student-Selected Studies (SSS) Biochemistry presentation slides: **websites**	47	9	Plagiarism was found in this slide presentation during a re-call of previous work by a student after a previous hearing – ie this was plagiarism not found by any marker. SafeAssign failed to detect extensive copying from an openly available website containing clinical education cases. The student was supposed to be reporting on details from their consultation with a real patient – an exercise in real clinical practice. Instead they copied a significant part of the clinical history details from this website (inlcuding liver test results). Turnitin detected this, and showed exactly which details had been copied. SafeAssign detected a match in one of the bibliography references which was spurious as it matched only the root part of the NLM URL to articles. This Level-2 offence would not have been detected at all by SafeAssign, but would have been revealed by Turnitin, had we been screening this courseork for matches.
3. MBBS UG: Analytical review (critical appraisal of a **medical research article**)	58	12	SafeAssign Picked up one part of the abstract of the paper being reviewed in a Website of abstracts of medical research articles. It did not detect the large amounts of that same article copied into the document, which were found by Turnitin. This is not suprising, given that the database of academic titles used by SafeAssign is one designed for researchers into Business and Commerce. SafeAssign delivered a warning that there were characters in this document which might indicate avoidance of detection. However it did not indicate where in the document these suspicious characters had been found, which is very unhelpful. There were several scientific/statistical characters in the article but not known if one of these triggered this warning.

			This case, assigned Level-2 plagiarism, would most likely have been assigned a Level-1 if SafeAssign had been relied upon for investigating the case.
4. PGR dissertation chapter 1**: websites, health-related published research articles** , **another thesis**, and many other sources Note There was no reference list in the original document – this formed a separate document in the dissertation.	51	8	Extensive copying from many different sources (inlcuding many readily available websites) detected by Turnitin was hardly detected by SafeAssign. Compare, for example, p.7 of the thesis which in the Turnitin report hardly has any <u>non</u>-highlighted text on the page, with the SafeAssign report (p.6 of the pdf) which has managed to detect only one in-text citation. This is typical of the entire document. Most SafeAssign matches are identified as from 'another student's paper' which is unhelpful.. An example of <u>patch-paraphrasing</u> of website material detected by Turnitin and not by SafeAssign is on p.5 of the Turnitin (source #25) and on p.3 of the SafeAssign report. Blackboard's claim that SafeAssign detects partial matches is seen <u>not</u> to be justified here and in many other instances. The student copied extensively from <u>another thesis</u> (Turnitin source #4)(p. 52ff.) and this was not detected at all by SafeAssign.
5. PGR Dissertation Chapter 2 - selective copying of published article to describe own results Note There was no reference list in the original document – this formed a separate document in the dissertation.	27	7	Copying of sources included websites and published articles. In p. 82 of the document the student has copied conclusions from another article and selectively substituted the name of a different condition , in order to help describe the results of their own systematic review. Turnitin shows clearly where the names of the conditions have been edited, as well as other minor changes such as substituting 'behaviour' for the original American spelling 'behavior'. The verbatim highlighting (and, importantly, lack thereof) of the copied text by Turnitin is crucial for determining the nature of plagiarism being committed, and also provides evidence of the intent to deceive. SafeAssign detected only in-text citations and section headers in this chapter. It did not find any text from other sources. An offence which would likely have been graded medium or high level using Turnitin would not even have been considered if SafeAssign had been relied upon.

Suitability for batch screening
SafeAssign's uninterpretable overall scores, and its abysmal detection capabilities make it unsuitable for 'batch' screening of coursework. We have shown here at least three cases (#2, #4 and #5) which would not have been detected by SafeAssign if used in screening mode, due to its failure to detect any significant matched text. If students use a software to self-check their work

which fails to detect sources, or parts of sources, which have been copied, this will give a false and a confusing impression about the acceptability of this kind of practice. Blackboard is infuriatingly agnostic about interpreting the scores and the matches in its helpsheets on SafeAssign, and provide no instruction on this. They are even careful not to refer to SafeAssign as 'text-matching' software.

Fitness for purpose
Overall I conclude that SafeAssign is not fit for purpose for investigation into, primary detection of, or education about, copied text in academic work. I believe that if the University considers this a sufficient tool for any of these purposes then its reputation and standards are at a very high risk. This would be pretty newsworthy. Also I believe that this would be displaying a disregard for the difficult and time-consuming work of plagiarism officers and many others in trying to uphold these standards. Any apparent cost-saving is false economy.

Report from BIO (Iain Barr)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
The five cases (some of which were close ca 30 000 words) was comparable. In this case the entire five searches in Safeassign took less than 2 minutes to return a result. Turnitin took 4.5 minutes to do the same. This was run in the afternoon at the suggested 'slow time' as we should be testing at any time of day. The uploading and results were fast and anyone used to using the UEA network will be surprised at how swift the results are returned. The creation of the PDF files was also swift, faster in Safeassign than turnitin but by a few seconds.
Ease of use
Both of these programs are easy to use, they are simple once you are through the odd blackboard front end at UEA and thereafter simple to use for all of the normal processes involved in these plagiarism cases.
Effectiveness
Student 1: Safeassign swifty recognised that this was 100% match to an item of coursework previously submitted. This remained 100% whether the comparison was with or without the figures and images that were the main subject of the plagiarism in this case. Turnitin did also recognise that this text had been submitted previously and flagged up a 100% match. Turnitin also recognised a 29% match to internet sources but did not provide detail of these sources. It also did not recognise the images as being taken from an internet source. In this case both software performed equally and would need staff input to identify the figures. **Recommendation either** The four other cases presented here (anonymised) were a range of cases whereby independent (of these two programs) had identified between 5% and 25% plagiarised material. This formed the basis of an interesting comparison. Each case was clearly identified by the assessor as being at least part plagiarised. Student 2: Safeassign: 12% and correctly picked up the majority (80%) of the plagiarised sections of the work. Turnitin 28% and picked up all of the plagiarised material. **Recommendation Turnitin** Student 3: Safeassign 0% match and failed to pick up anything. Turnitin 5% match and correctly identified the plagiarised sections. This was from wikipedia (4%)- which is especially concerning that safeassign failed to identify this as a source. **Recommendation Turnitin** Student 4: Safeassign 3% match but all of these were the references. Turnitin 21% match of which 5% is the actual plagiarised sections- all correctly identified. In this case turnitin did also highlight the majority of the reference section as a match. The outcome if relying on the software would clearly be cause for concern here: **Recommendation Turnitin** Student 5: Safeassign 2% match all of this is in the references section and so would not be useful in a plagiarism case. Turnitin 18% match and correctly identified 4 main internet sources used and copied in this report. All of the references identified were flagged with a <1% marker and so could simply be discounted on brief visual inspection. **Recommendation Turnitin** The test resulted in four recommendations for Turnitin and the difference between Turnitin and Safeassign are so large and concerning that they would have lead to a different outcome in each of these four cases, should the software have been used in isolation. If Safeassign alone was used; Three would have resulted in a no plagiarism verdict, where there was plagiarised material

(false negative) and one would have lead to a verdict of low whereas it should have been medium (as correctly identified by Turnitin).

The first case (student1) would have had no difference between the software, with both picking up the collusion elements but neither picking up the plagiarised diagrams and images.

Matches

The matches made by safeassgin were by and large references and not bodies of text that we should be concerned about in academic plagiarism. The inability of Safeassign to match large sections of text from wikipedia in the case of Student 3 is very worrying. Turnitin performed much better and would lend a degree of scalability whilst retaining confidence that safeassign would not.

Suitability for batch screening

Both of these programs would have been suitable for batch screening. However using Safeassign you would need to expend around 4-6 minutes per script to run through each Safeassign report and discount the reference matches. In a typical class of 80 students this would mean over 6 hours additional work. Most of this would be meaningless as the values of match overinflated by the references which would be picked up as a 3-5% match on every script- a figure well within the typical for a low level offense.

The screening of Safeassign reports would not be compatible with shortening the turnaround time for coursework return that UEA is demanding of staff.

The lack of identifying commonplace sources in Safeassign is very concerning indeed and would not lend itself to confident batch screening.

I would be uncomfortable recommending to students that they pre-screen their work in Safeassign as they would play the minimum match game which would be much simpler to avoid in Safeassign than Turnitin

The use of Turnitin for batch screening would allow a degree of certainty in the overall match metric so that staff can target time on the high level and select some low level matches to further scrutinise.

I would recommend that if Turnitin is used that a staff member at the UEA repeat this sort of test every 2 months to ensure that Turnitin is still maintaining performance at this level. UEA needs to ensure that it is always maintaining this level of confidence in the program. Also i would recommend that anti-software matching software be investigated.

Fitness for purpose

It is very clear that Safeassign is not fit for purpose. In my mind a program of this type that fails to identify Wikipedia sources is not worth using and can potentially lead to false confidence on the part of students in their pre submission checking.

Staff time: time spent checking the Safeassign reports for referencing matches is wasted time. Approximately 55% of my time in assessing the Safeassign reports was wasted on considering and rejecting references and bibliography. 40% of my time was spent on assessing the matches of broken sentences that are clearly not under suspicion of plagiarism (paraphrasing maybe but not plagiarism). The last 5% was spent in assessing genuine matches. This is a balance that puts the onus on staff to discriminate, which we can do without this program- it is not an aid but an encumbrance to the same end result.

I would fear the use of the Safeassign program in batch screening as the staff time devoted to checking references would outweigh the weighting put on references in SCI. Turnitin in contrast seems to do a reasonable job in finding sources of plagiarised material. The only shortfall was in student 1 where the plagiarised diagrams and figures were not highlighted.

I would also add that the PDF's produced by both software programs have issues and i would want a different quality from either should they be used. Safeassign produces long and badly formatted PDF's and Turnitin produces PDFs with low resolution. I would recommend that

should Turnitin be used that a higher quality PDF be required on the UEA print to PDF default- a simple higher quality plugin may be required.

Report from LDC (Cecilia Rossi)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
<u>**CRITERIA**</u>
Speed
From my experience, I don't see much difference between the two TMS systems regarding the speed at which the originality reports are generated.
Ease of use
I have trained to use both systems and must say that I found the process of uploading essays, generating the reports, saving them as pdf files much simpler with Turnitin than SafeAssign. When it comes to interpreting the originality report, there is no doubt that the reports generated by Turnitin are overall more user-friendly, and, above all, easier to understand for the students when they are summoned to a hearing.
Effectiveness
Essay 1: This piece of formative work was brought to my attention by the module convenor, but it was spotted by one of our associate tutors, who only noted in the script the sources she had found, namely, www.gradesaver.com and www.enotes.com I ran the assignment through SafeAssign (as it was the only TMS system available then, in March 2016) and used the report for the hearing, which I knew would be effective enough because it had picked up these two sources. The hearing proceeded with no problems and the outcome was Low Level. The Turnitin report that I generated recently shows a lower match (29% as opposed to 37%) but it is more reliable as it has picked up exactly the words that have been lifted from these sources, and not the surrounding words (please see reports attached). **Essay 2:** This is a summative piece of coursework which I recalled: it had been marked (awarded a 65) but I recalled it for a hearing with a student whose work previously had been established to have been largely plagiarised. The student never came to the meetings, and has subsequently withdrawn from the university. At the time (early Spring semester) I only had the SafeAssign report which shows a match of 27% as opposed to the Turnitin report which shows a match of 70%. The latter is much more accurate and in keeping with the other pieces of coursework which the student submitted. So the SafeAssign report would not have been effective had we needed to establish the exact amount of plagiarised text. **Essay 3:** This is a Level 6 Literature Dissertation brought to my attention by the two markers who listed a series of possible sources, including the Wikepedia entry which the Turnitin report picked up, but the SafeAssign TMS did not. The Turnitin report was used in the hearing and the outcome was Low Level as there were extenuating circumstances brought up by the student and which we accepted in light of the evidence. The SafeAssign report, which only picked up a 3% match, would have been completely ineffective at said hearing. **Essay 4:** This was a case from three years ago. The Turnitin report was instrumental in proving the student had been lifting the odd phrase from here and there using several online sources. The overall match with the Turnitin report was 27% - as opposed to 1% in the SafeAssign report. As in the previous case, the latter report would have been of no use in the hearing. **Essay 5:** Another case from a Level 6 student, and from a few years ago. The Turnitin report was used to great effect as it pointed me in the direction of the sources the student had used. As in the third case, the Wikepedia entry used by the student and not acknowledged was picked up by the Turnitin report but not by SafeAssign.
Matches
Turnitin   SafeAssign Case 1: 29%   37% Case 2: 70%   27% Case 3: 16%   3% Case 4: 27%   1% Case 5: 32%   12%
Suitability for batch screening

My fear is that, given the discrepancy in the similarity index, the use of Safe Assign will give students a sense of false security. Also, it does not pick up Wikepedia entries, which, at least in LDC, are commonly used.

Fitness for purpose
From my answers in 'Effectiveness' it is clear that I must conclude that Turnitin is overall much more effective and therefore more reliable.

Report from Jennifer Bowler (PSY)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
Report PSY UG 1 Submitted Turnitin 10:59am. A few seconds to upload; one to two minutes to generate originality report. Submitted SafeAssign 4:02pm. Originality report generated within two minutes. Report PSY UG 2 Submitted Turnitin 10:57am. A few seconds to upload; one to two minutes to generate originality report. Submitted SafeAssign 3:57pm. Originality report generated within two to three minutes. Report PSY UG 3 Submitted Turnitin 1:59pm. A few seconds to upload; two minutes to generate originality report. Submitted SafeAssign 7:56pm. Originality report generated within three minutes. Report PSY UG 4 Submitted Turnitin 6:10pm. A few seconds to upload; two minutes to generate originality report. Submitted SafeAssign 8:07pm. Originality report generated within two to three minutes. Report PSY UG 5 Submitted Turnitin 6:25pm. A few seconds to upload; one minute to generate originality report. Submitted SafeAssign 8:07pm. Originality report generated within two to three minutes. Generally, I have found the speed of SafeAssign and Turnitin to be comparable in terms of actually generating the originality report, although the former (SafeAssign) is considerably more 'clunky' to use and it can take longer overall to process an assignment for this reason.
Ease of use
Report PSY UG 1 Functionality of SafeAssign is much less intuitive than Turnitin. Assignment has lost its formatting within the SafeAssign originality report, making it less clear to read and interpret, whereas this was not the case in Turnitin. Several more steps to save the SafeAssign report as a pdf. Final saved report is not colour coded, impeding straightforward interpretation relative to Turnitin, although comparison boxes of original versus matched source text are provided at the end of the SafeAssign report. Report PSY UG 2; PSY UG 3; PSY UG 4; PSY UG 5 Turnitin: it was easy to exclude sources as necessary and to edit settings (e.g. number of words/exclude bibliography) of the originality report, increasing its usefulness as a guide for decision-making. Turnitin and SafeAssign: it would be useful to also be able to exclude appendices.
Effectiveness
Note that all Turnitin reports were filtered to exclude matches of less than 4 words, whereas this was not possible using SafeAssign. Below I report the matched text percentage guides generated by each report and comment on the overall effectiveness of each software in correctly and usefully identifying matches throughout each assignment. Report PSY UG 1 Turnitin 28% including reference list, 26% excluding reference list SafeAssign 10% with reference list but no word restriction (cannot see how to exclude bibliography/apply word restriction as did for Turnitin). The matches that it has detected are irrelevant (within reference list/appendices only) and it appears to have missed the relevant matches (which were predominantly in the introduction, as detected by the marker and corroborated by evidence in the form of the source journal articles). In contrast, the plagiarised

content in the introduction was usefully highlighted by Turnitin. This is a significant omission by SafeAssign and effectively renders it redundant in terms of its contribution to the evaluation and decision making process.

Report PSY UG 2

Turnitin 37% including reference list, 27% excluding reference list – appears to have automatically excluded appendices (student labelled as 'Annex') when 'Exclude bibliography' selected.

SafeAssign 40% (including reference list and appendices)

The main problem area of this report was the introduction. Whilst some of the unreferenced/inappropriately cited journal article sources within the introduction were correctly detected by both SafeAssign and Turnitin, Turnitin was considerably more sensitive and detected plagiarised sentences that were omitted by SafeAssign. The SafeAssign percentage was artificially inflated by its inclusion of more content from the reference list and appendices.

Report PSY UG 3

Turnitin 6% (excluding reference list)

SafeAssign 5% (including reference list, so fewer relevant matches detected overall; those sources matched in text are repeatedly "another student's paper" – more information would be useful here).

Report PSY UG 4

Turnitin 38% (excluding reference list)

SafeAssign 5% (including reference list). Parts of the originality report are indecipherable. Appears to have missed a number of journal articles that Turnitin detected.

Report PSY UG 5

Turnitin 45% (excluding reference list)

SafeAssign 11% (including reference list). Matches are all other student papers; appears to have missed large sections of plagiarised text in the introduction and failed to detect journal article sources detected by Turnitin. Formatting of the SafeAssign report is also very difficult to use effectively, reaching 104 pages, whereas the Turnitin report is appropriately formatted (40 pages).

Matches
Overall, Turnitin is more reliable and sensitive at detecting plagiarised content, frequently detecting sentences from journal articles that were omitted by SafeAssign, as well as from other sources (e.g. student papers). Turnitin also provides more information on the other sources, which is useful when checking and interpreting the highlighted content. The extent of the omissions by SafeAssign within this evaluation process means that it would not always be possible to use the report effectively when evaluating a case, and decisions would need to be based on manual text screening and collated evidence, not the report itself. In contrast, Turnitin has successfully highlighted substantial portions of plagiarised content that were omitted by SafeAssign.

Suitability for batch screening
I have not been required to batch screen assignments and so cannot comment on the suitability of either software for this purpose.

Fitness for purpose
Turnitin is undoubtedly more useful that SafeAssign in evaluating cases. SafeAssign has repeatedly failed to detect content that has been manually identified as plagiarised (i.e. by the marker/POs as part of the evaluation process), whereas Turnitin has successfully highlighted this content. In addition to detecting duplicated material more reliably, apparently drawing on a larger range of journal articles, Turnitin also has much more intuitive and applicable filters that make it more sensitive and specific, and hence meaningful to the appraisal process. The presentation of the Turnitin report is also much superior to that generated by SafeAssign, in which assignment formatting is frequently corrupted (as for instance with reports 4 and 5 of this evaluation), and hence I find the Turnitin report easier to use and more helpful when processing cases.

Report from ENV (TYN) (Rachel Warren)

Comparative Evaluation of SafeAssign and Turnitin, Summer 2016
CRITERIA
Speed
Both Turnitin and SafeAssign ran promptly during my tests with each taking less than a minute to run a single essay of about 4000 words. However because the SafeAssign results are less clearly presented it takes the PO longer to come up with an assessment of the amount of plagiarism that has really been detected by the software.
Ease of use
Both Turnitin and SafeAssign are easy to initiate in terms of uploading a student's work and setting it going. However, the way that SafeAssign presents its outputs is much less clear and it more difficult to interpret the amount of plagiarism that has really occurred.
Effectiveness
SafeAssign fails to detect blatant plagiarism committed by 7 of the 8 cases I reviewed below. Only for 1 case did SafeAssign come up with anything like the same amount of matches. I outline some particularly poignant examples below: **Case 2:** This was a student's chapter from a piece of group work and the marker had detected that much of it was taken from a single internet source. Turnitin picked up the sections that the marker has already shown me that matched an internet site, and it also identified several other shorter examples of plagiarised text not located by the marker. The discussion in the plagiarism meeting was usefully backed up by The turnitin report, as the student was questioning whether it was really possible to reword some of her sentences. It was also very useful to the marker in being able to identify how much of the work was really student's own when it came to reassigning a new mark (as this case had been classified as medium level). SafeAssign detected almost no plagiarism at all in this document.   **Case 6:** This was the case of a resubmitted dissertation. This student was known to have had a lot of difficulties with his work, and unfortunately the marker also detected plagiarism in his submitted work. He highlighted several sections of text which were obviously in much better English than that reported by the student, and had found one of the internet sites where some of the text had come from. I ran the dissertation through Turnitin and was able to identify this internet source and also several others. Turnitin correctly identified the passages where the student had explained his own results, as lifted from others' writings. After the meeting the offence was classified as medium and the marker was asked to remark the work. With Turnitin having shown him that even less of the work was the student's own, he was able to remark the script appropriately. SafeAssign, on the other hand, detected only a low level of plagiarism in the work and failed to identify the crucial passage where the main crux of the student's supposed discussion of his own results had actually been taken from elsewhere.   **Case 5:** In this case, a marker identified plagiarism and because of the particular circumstances I was asked to check some previous work that the student had submitted as well as the one in question. Turnitin enabled me to identify that this student had committed serial plagiarism. It turned out the meeting that the student was having a number of problems that he needed help with. However, SafeAssign detected some plagiarism in the current piece of work and almost none in the previously submitted ones. So without Turnitin I would not have found out that this student was a serial plagiariser.
Matches

Assignment	SafeAssign	Turnitin
1	7%	56%
2	1%	47%
3	5%	62%
4	11%	27%
5	15%	35%

6	7%	38%
7	4%	13%
8	60%	68%

Suitability for batch screening

In view of dismal performance of SafeAssign in my tests above, and also in the tests performed by Alicia in her ENV batch screening exercise, SafeAssign is completely unsuitable for batch screening. It is essential that we have a consistently reliable tool to identify plagiarism in student work. An unreliable tool allows academic standards to slip, as POs become much less able to correctly identify levels of plagiarism and to defend their judgements in meetings. It also creates potential bias as SafeAssign can detect plagiarism in some pieces of work (eg example 8 above). For example in case 3 above a guilty student would potentially have faced no penalty under SafeAssign whereas in case 8 they would, and in fact Turnitin identifies very similar levels of plagiarism in both cases.

Fitness for purpose

SafeAssign is clearly unfit for purpose. Its outputs are also difficult to interpret. If we were to continue using SafeAssign, a PO would not be able to easily confirm suspected plagiarism to a marker without manual checking. Then, either the PO or the marker would have to manually search the work in order to do the job properly. This would be enormously time consuming. The likely outcome is that people would simply not have time to do this very often, and even if they did, how to ensure that manual checking is fair? It would depend on exactly which phrases were typed into which search engine.   On the other hand Turnitin is an excellent consistent tool for identifying plagiarism, useful in my work with markers where plagiarism is first identified, and also in preparing reports ahead of meetings, where the reports are essential in order for the POs to correctly assess the amount and nature of the plagiarism. Without this tool, PO's ability to enforce the university's policies would be largely eroded.