

# **Either my students are getting naughtier, or the tools are getting better!**

---

**Fintan Culwin**

*London South Bank University, UK*

Two empirical studies of non-originality in student work are presented. The first includes four years of data from approx 650 first year students. The second study contains three years of data from approx 800 final year students. Despite methodological problems with the tools used the data suggests that amount of non-original material used by the students remains constant; but there is some evidence that the way final year students had used the material changed.

## **Introduction**

Software tools to assist with the detection of non-originality are well established but variable in their efficacy (Satterwhite & Gerein 2001). This paper will report upon the rates of non-originality detected by two different tools over three sequential cohorts of computing students at London South Bank University (LSBU), measured at the start and end of their undergraduate progression. The tool used for the first year students was OrCheck which makes programmatic use of the Google search engine (Culwin & Lancaster 2004). The tool used for the final year students was the JISC supplied Turnitin service which makes use of its own private detection engine technology. The use of the tools is only a part of a wider and more comprehensive pro-active academic misconduct policy (Carroll 2002, Culwin & Lancaster 2001)] in the computing department at LSBU.

Further discussion on the methodologies involved in collating evidence regarding the extent of academic misconduct. Followed by the pro-active context in the department where the data was collected. Essentially students are actively introduced to issues of misconduct at the start of the first year, as described in Culwin (2006); reminded in a core unit at the start of the second year and again at the first project lecture at the start of the final year. This is complemented by publicised Intranet resources and advice and also by wall poster displays.

## **The first year studies**

Precise details of how the data was collected. Essentially as reported in the first conference and subsequently in Culwin (2006); first year students in the first six weeks were required to submit a 1000 word essay on an individualized topic from the history of computing. These essays were then subject to non-originality investigation using the Google powered OrCheck system. The resulting measures of non-originality were then collected and collated . . . .

---

## **Corresponding Author**

Fintin Culwin, Faculty of Business, Computing and Information Management,  
London South Bank University, Borough Road, London, UK

**Email:** [fintin@lsbu.ac.uk](mailto:fintin@lsbu.ac.uk) **Tel:** +44 (0) 207 815 7434

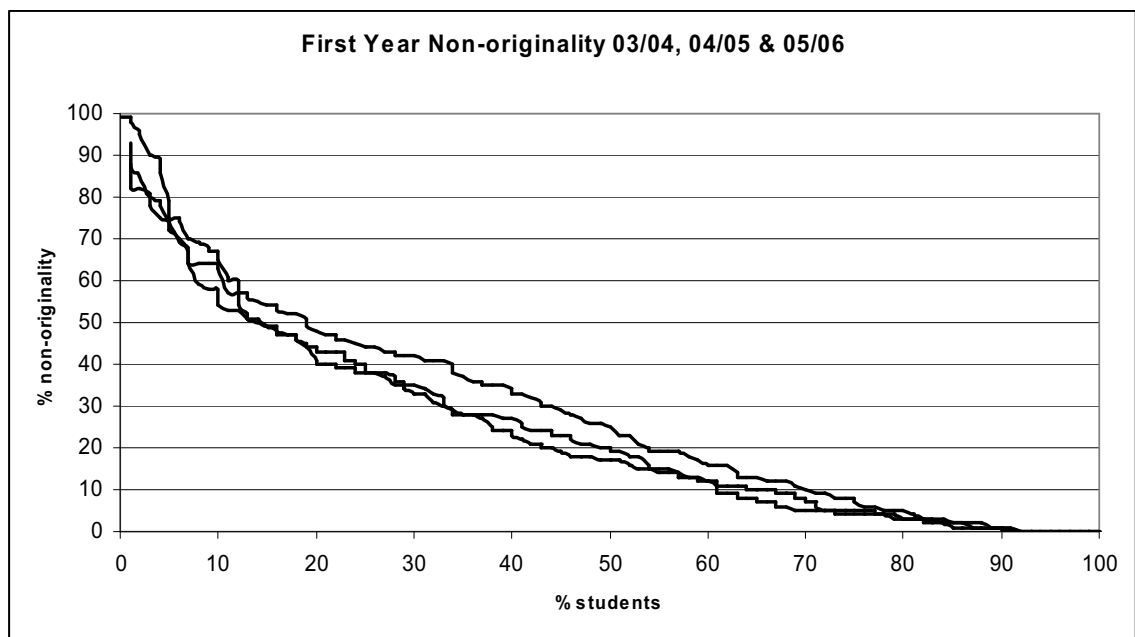
---

Year	n	Av	Std
02/03	156	17.2	19.2
03/04	207	24.8	24.7
04/05	147	23.8	22.4
05/06	134	28.0	23.5

**Table 1:** First Year Summary Results

Table 1 shows the summary results for the four years. The columns are the academic year of the cohort, the number of students in the sample, the average percentage non-originality score and the standard deviation. This latter statistic is a measure of how widely distributed the values are. To have a standard deviation so large compared with the average indicates that there is a wide range of behavior.

The first year of the investigation only included computing students and used software that was still under development. Accordingly the measures are not directly comparable with the succeeding years. For the remaining three years there was no essential difference in the software, and statistical tests<sup>1</sup> indicated that there was no significant difference between the three years.



**Figure 1:** First Year Non-Originality 03/04, 04/05 & 05/06

Figure 1 shows the pattern of non-originality for each of the three years. This graph also emphasises the similarity between the three years. The information is shown as cumulative non-originality. For example and considering only the upper line, it shows that 20% of the students submitted work that was shown to be more than 50% non-original. Or reading in the other direction and considering only the middle line, 50% of students submitted work that was shown to contain less than 20% non-original content.

<sup>1</sup> Three repeated two tailed t-tests:  
05/06 cf 04/05, 05/06 cf 03/04 & 04/05 cf 03/04 all non significant even at the 0.1 level.

In the mid-range of the graph, between approx 20% and 70% of %students, the lowest line is 03/04, the middle line 04/06 and the upper line 05/06. As lines which are nearer to the axes are indicative of less non-originality this gives some indication of a possible trend towards more non-originality.

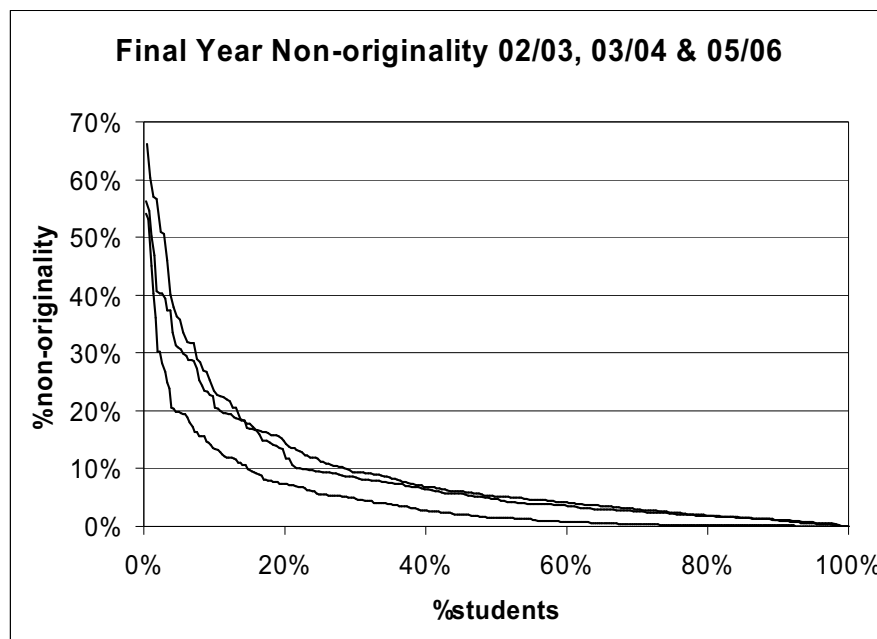
### The final year studies

Precise details of how the data was collected. Essentially every final year project was collected and analyzed by TurnItIn via the JISC service.

Year	n	Av	Std
02/03	306	4.7	7.9
03/04	267	8.9	10.1
04/05	213	9.0	11.9

**Table 2:** Final Year Summary Results

Table 2 shows the summary results for the three years; the values are as described for Table 1 above. Statistical tests showed very significant differences between the 02/03 data and both of the other years. There was no significant difference between 03/04 and 04/05.



**Figure 2:** Final Year Non-Originality 03/04, 04/05 & 05/06

Figure 2 shows the data as a graph and illustrates the similarity between the two latter years which are the upper two lines; and their collective dissimilarity with the first year, shown on the lower line. The cumulative non-originality can be interpreted as for the first year graphs in Figure 1 above.

<sup>2</sup> Three repeated two tailed t-tests:  
02/03 cf 03/04  $p < 0.05$ , 02/03 cf 04/05  $p < 0.05$ , 03/04 cf 04/05  $p > 0.5$ .

## Discussion & conclusions

Comment upon the differences between the first year and final year gross measurements, together with the reasons and implications.

Reminder upon the reasons for the anomalous 02/03 first year data. Consideration of the possible reasons for the anomalous 02/03 final year data, (if anything the values should have been higher that year.

Comment upon the strange consistency of the remaining data.

Enumerate the possible explanations:

- Real outcome – tools are accurately measuring stable behaviour.
- Students are getting naughtier but tools are getting less effective.
- Students are getting less naughty but tools are getting better.

Introduce more anecdotal qualitative evidence. The manner in which non-original material is being used in final year projects has changed. Students are using it more responsibly and this is evidenced by fewer numbers of projects being subject to formal investigation.

Overall conclusion – unknown as per the title & the need for more studies based upon artefacts of actual behaviour rather than questionnaire studies. However the pro-active policy in the department has changed the attitude towards non-original material.

## References

Carroll J. *A Handbook for Deterring Plagiarism in Higher Education*, Oxford Brookes University, ISBN 1-873567-56-0

Culwin F & Lancaster T (2001), *Plagiarism Issues for Higher Education*, *Vine* 31(2), 36-41

Culwin F & Lancaster T (2004), *Beyond Learning Objects: Towards Learning Beans*, *Proc. ASCILITE 2004*, pp 230-239, ISBN 0975170236, UWA Perth Dec 2004

Culwin F (2006), *An active introduction to academic misconduct and the measured demographics of misconduct*, *Assessment & Evaluation in Higher Education*, 31( 2) 167 - 182

Satterwhite R & Gerein M (2001) *Downloading detectives: searching for on-line plagiarism*, available at <http://www.coloradocollege.edu/library/course/webplag2.html> (accessed 20 Feb 2006).