**University Research Ethics**

**APPLICATION REFERENCE:**

**for office use only**

**Committee**

**Application Form**

**Checklist**

|  |  |
| --- | --- |
| **Name of applicant: Andrew Wicks** | |
| **Faculty/Directorate: ACH/CIS** | |
| **Title of research: Content factors which influence the success of a VLE. – Part 1** | |
| **These papers must be attached to this application form (please tick):** | |
| * Participant information sheet | N/A |
| * Participant consent form | N/A |
| **These papers may be required (tick if included):** | |
| * Letters (to participants, parents/guardians, participating institutions etc.) |  |
| * Questionnaire(s) or indicative questions for interviews |  |
| * Advertisement /flyer/copy of message inviting participation |  |
| * [Annex I](http://www.gre.ac.uk/__data/assets/word_doc/0009/644859/UREC-application-form-annex-1.doc) - Drugs and medical devices |  |
| * [Annex II](http://www.gre.ac.uk/__data/assets/word_doc/0010/644860/UREC-application-form-annex-2.doc) - Research involving the storage of human tissue |  |
| * [Annex III](http://www.gre.ac.uk/__data/assets/word_doc/0011/644861/UREC-application-form-annex-3.doc) - Ionising radiation |  |

**Has the form been signed? YES / NO**

**Have any annexes been signed where necessary? YES / NO**

Revised October 2015

### **SECTION 1: APPLICANT DETAILS**

|  |
| --- |
| * 1. Surname Wicks Forename Andrew Title Mr. |
| Faculty/Directorate: ACH/CIS |
| University address, including Faculty Department  University of Greenwich, ACH/CIS, Room QM420, Old Royal Naval College, 30 Park Row, London SE10 9LS |
| University Tel.: 020 8331 8717 E-mail: wa02@gre.ac.uk |
| * 1. Are you a student?   A member of staff?  A member of staff applying as a student? **Yes**  Other? |
| Programme of study (if applicable)  **MPhil/PhD** / EdD / Masters by Research / MSc / MA / BSc / BA / DipHE / other (please specify) |
| If you are a higher degree student, has your research project been approved by your Faculty Research Degrees Committee? Indicate YES or NO. **YES** |
| If YES, when? **10th Jun 2015** What is the FRDC reference number? **FRDCACH/14/M-5/4.5**  If NO, why not? |
| * 1. What is the primary purpose of the research? (Please indicate YES or NO) * Educational qualification **YES** * Internally funded research **YES**      * Externally funded research (please provide details of funding)      * Other (please specify)………………………………………………………………............................ |
| * 1. Project supervision – give the name of the research supervisor(s) and their contact information * Prof. Liz Bacon, University of Greenwich, QM353, Old Royal Naval College, 30 Park Row, London SE10 9LS (email: [e.bacon@gre.ac.uk](mailto:e.bacon@gre.ac.uk), Ext: 8544) * Prof. Lachlan MacKinnon, University of Greenwich, QM353, Old Royal Naval College, 30 Park Row, London SE10 9LS (email: [l.mackinnon@gre.ac.uk](mailto:l.mackinnon@gre.ac.uk), Ext: 8559) * Dr. Cos Ierotheou, University of Greenwich, QM334, Old Royal Naval College, 30 Park Row, London SE10 9LS (email: [c.ierotheou@greenwich.ac.uk](mailto:c.ierotheou@greenwich.ac.uk), Ext: 8655) |
| **1.5**Details of any other co-researchers within the university  None |

|  |
| --- |
| * 1. Details of any other co-researchers external to the university   None |
| **1.7** Membership of professional bodies - are you or any co-researcher(s) a member of any professional, or other, bodies which set (i.e. require compliance with) ethical standards of behaviour or practice such as the British Psychological Society, Nursing and Midwifery Council, medical Royal Colleges etc.? If so, please specify.  **None** |

### **SECTION 2: PROJECT DETAILS**

|  |
| --- |
| * 1. What are the principal research questions posed by this research? Describe briefly, in lay terms, the proposed research project including step by step methodology, and its potential outcomes and benefits (no more than 250 words).   **Research question**: Can a statistical technique be developed which is robust enough to gauge whether a new teaching technique is effective?  **Potential outcomes**: Studies in education have a problem of comparing like with like. If a new technique is being trialled (in this case a new type of e-learning environment), then how can the effectiveness be measured? Allowing part of a group to use the new technique would be impractical since it is web based and can, therefore, be accessed. It would also be unethical because one section of students could have an advantage over another (in whichever direction).  Comparing year-on-year for that course alone would not work because the academic, social and work ethic mix varies year-on-year.  What is needed is a technique which compares the relative position of one course to the others in the same year. If the new technique is effective, then the relative position of the course when compared to others should shift.  **Potential benefits**: This study seeks to create a statistical measure, called learning dissonance, to gauge the probability that this change was due to the new technique (or any other technique).  **Methodology**: If students were homogeneous items then one would expect that their scores in different courses would be the same, i.e. if a student achieved 56 in subject X then they would also achieve 56 in subject Y. The average differences between their scores on each of their courses is the learning dissonance for that student. A geometric mean of these learning dissonances between two specific courses is the learning dissonance score for that pairing in a particular year. The difference in learning dissonance between successive years would be a measure of the effectiveness/ineffectiveness of the technique being investigated. For example, if the course being investigated increases its learning dissonance then the technique would have had a positive effect on the scores of the students. |
| **2.2** Are any of the following involved? (Please indicate YES or NO)   * Intrusive procedure e.g. questionnaire, interview, focus group, diary **NO**   (attach a copy of your questionnaire or indicative questions) **N/A**   * Invasive procedure e.g. venepuncture, tissue sampling **NO** * Physical contact **NO** * Covert observation or covert filming / recording **NO** * Children / young people (under 18) – please include age of participants **NO** * Vulnerable people (elderly, physically or mentally ill, people with learning difficulties, in care, bereaved, prisoners, other) **NO** * Drugs, medicinal products or medical devices (if YES, complete [Annex I](http://www.gre.ac.uk/__data/assets/word_doc/0009/644859/UREC-application-form-annex-1.doc)) **NO** |

|  |
| --- |
| **2.3** Has there been a pilot study for this research? (If YES, please give details)  Yes – COMP1148 was compared to the other first year computing courses to assess the practicalities of this approach. This analysis is shown in Appendix A. |
| **2.4** What is the proposed start date (i) of the project and (ii) of the fieldwork (if different)?  **Start Date**: As soon as approval is given.  What is the proposed end date (i) of the project and (ii) of the fieldwork (if different)?  The end date would depend upon the date of submission of the dissertation. Currently, this would be 2018/19. |

### **SECTION 3: PARTICIPATION AND CONSENT**

|  |
| --- |
| **3.1** What are the selection criteria for the proposed participants in the study?  All students who have taken a course in the department of Computing and Information Systems from the 2010/11 academic year onwards. |
| **3.2** How many participants are to take part?  Approximately, 300+ students per academic year (although, many will reoccur in several years as they progress through the university). |
| **3.3** How will prospective participants be recruited / contacted and informed about their role in the project? (Give details and attach your participant information sheet, advertisement, email etc.)  There will not be any active participation. Their course scores will be obtained from the university BannerWeb system for the years 2010/11 to the academic year in which the study ends (either 2017/18 or 2018/19). PAS has agreed to provide the data subject to ethical approval. |
| **3.4** Where will the interaction with participants take place? e.g. online, classroom, public facility, laboratory, office, home etc.  There will not be any active participation. |
| **3.5** Are any external bodies’ premises or resources to be used? Please indicate YES or NO and give details of permission sought.  No. |
| **3.6** What is the expected total duration of participation in the study for each participant? e.g. 20 minutes to complete a questionnaire, an hour for an interview, etc.  There will not be any active participation. |
| **3.7** Is consent to be obtained using the UREC consent template? (Please indicate YES or NO and attach your [consent form](http://www.gre.ac.uk/__data/assets/word_doc/0003/644862/UREC-Consent-Form-final.doc)). If NO please indicate how consent is to be obtained, and attach a copy.  No. |
| **3.8** If children or young people (under 18) are involved, please say how consent will be sought, from both the children / young people and their parents, guardians or those acting *in loco parentis* (e.g. school).  Not applicable. |
| **3.9** Will any payment, incentive or reimbursement of expenses be made? (Please indicate YES or NO and give details, including amount)  No, as there will be no active participation |

### **SECTION 4: ETHICAL CONSIDERATIONS**

|  |
| --- |
| **4.1** What do you consider are the main ethical issues and risks that may arise in this research? (Refer to the Guidance on Ethical Approval for Research, [point 3.1](http://www.gre.ac.uk/research/ethics/what-might-be-the-ethical-issues-and-risks-that-arise-in-my-research)). What steps will be taken to address each issue?  **Potentially sensitive data**: The research will require the names and enrolment numbers of students, so that their scores in various courses can be analysed. However, the enrolment number will be encrypted using MD5 which will be used as the student ID when the data is analysed and published. This will maintain student anonymity whilst retaining the ability to see how they performed.  **Matters around the subject matter**: Student grades are involved, which will be sensitive to some. This issue is dealt with in the discussion below.  **Matters around participants**: There are no issues around vulnerability. All the records are for those who have completed courses in the Computing and Information Systems department of this university. Whilst some of these students will have been taught by me and all of those teaching these courses will be my colleagues, the information being used is historical and available to any member of the teaching staff via BannerWeb. The measures used to protect the identities of both staff and student are outlined below.  **Matters around the researcher**: There are no conflicts of interest in this research. The research seeks to develop a statistical measure of changes in the educational environment.  **Matters around the recruitment of participants**: Participants are not being recruited. The results used are available to all teaching staff of this university.  **The nature of the participants**: Only scores of students are being gathered. No interviews will be undertaken for this research.  **Risks or hazards to participants or researchers**: Participants will not be used directly in this research. There are therefore no hazards for them. There are no hazards for the researcher either.  **Location of participation**: There are no location related hazards for this study.  **Confidentiality issues**: The only confidentiality issues surround the harvesting of grades from BannerWeb. The identities of the students are needed to track their scores across courses. Each student will be assigned a random number as they enter the system. The identity information can be deleted once the study has been completed, leaving the random number to identify a student. The random numbers will not allow any tracking back to a real person, since the numbers will not be any order.  The same process will be used for lecturers. Their names will be needed initially for tracking, but can then be deleted, again leaving just their random number, once the study is completed.  Please see Appendix B for the details.  **Other health and safety issues**: There are no other health and safety issues. |

|  |
| --- |
| **4.2** Will personal data, as defined by the Data Protection Act 1998, be collected during the research (Refer to the Guidance on Ethical Approval for Research, [point 3.2](http://www.gre.ac.uk/research/ethics/am-i-collecting-personal-data-in-research))? Indicate YES or NO. If YES, give details of how you will deal with that data. **YES**  **What personal information will be collected?**: The name and enrolment number of students. The name of staff members teaching a course. However, the enrolment number will only be available in its encrypted form.  **Gain consent for collecting and processing the data**: No consent is being sought because the data is historical and held in BannerWeb on the university system. It would also be impractical to contact everyone who ever took a course. This is unlikely to be an issue since individuals will not be identifiable in the data sets once the study is complete and the findings released.  **Provide enough information about the project for the participant to be able to give informed consent**: No such permission is being sought because the data is available to any member of the teaching staff via BannerWeb.  **Security of the data**: The original, unanonymised data will only be held on the university system and a remote desktop will be used for ease of analysis. The university file system is secured behind a firewall, on password encrypted servers. Any work completed on this project outside the university would use the fully anonymised data as shown in Appendix B. Only fully anonymised data will be stored external to the university.  **Retention of the data**: The confidential data (names and enrolment numbers) would only be kept on the University drive until the study was completed, i.e. at the end of the Ph.D. which is expected to be in 2018/19.  **Disposal of data**: It is envisaged that the anonymised data will be kept in the repository of research data at the University of Greenwich so that other researchers can examine the data and the techniques developed. All identifying information will be removed from this version, as described above.  **Confidentiality of data**: There are no direct participants in this study. The personal data collected will be in separate text files so that it can be separated from the statistical data easily once the study is completed.  **Ensure that data is published only in anonymised form**: Please see Appendix B which show the layout of the data before and after release. As can be seen, all identifying information has been removed from the proposed release version.  **Future use of data**: The purpose of this study is to assess the practicality of the learning dissonance statistic. There is no need to collect any information irrelevant to this aim, e.g. names, enrolment numbers, names of courses or lectures teaching those courses. New approval would be sought if the study were to be extended, e.g. by the university for monitoring purposes. |

### **SECTION 5: INSURANCE AND FINANCIAL INTERESTS**

|  |
| --- |
| **5.1** Will this research be covered by an insurance policy (such as your own professional indemnity insurance) other than the University of Greenwich public liability policy? If so, give details.  **NO** |
| **5.2** Indicate by “YES” or by ticking one of the statements below:   * I declare there is no financial or other direct interest to me or my Faculty or Directorate arising from this study **YES** * I declare there is a financial or other direct interest to me or my Faculty or Directorate arising from this study (supply details) |

### **Signatures**

|  |
| --- |
| I undertake to carry out research in accordance with the University’s [Research Ethics Policy](http://www.gre.ac.uk/research/rec/rep). In the case of a higher degree, I confirm that approval has been given by the Faculty Research Degrees Committee. |
| **Signature of applicant Date**  **Print name** |
| I have discussed the project with the applicant, I confirm that all participants are suitably qualified to undertake this research and I approve it.  **Signature of supervisor (to be signed if applicant is a student) Date**  **Print name** |
| I have reviewed the project with the applicant, or applicant’s supervisor, and I confirm that all participants are suitably qualified to undertake this research and I approve it.  **Signature of** [**UREC representative**](http://www.gre.ac.uk/research/rec/faculty-representatives) **Date**  **Print name** |

### **Appendix A – The trial of calculating the learning dissonance statistic**

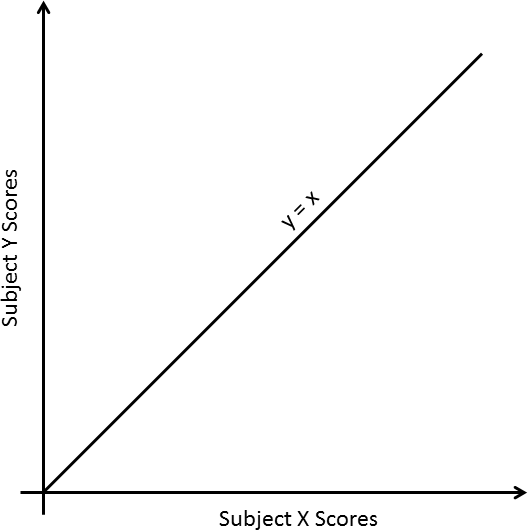
**Introduction**

The purpose of the following technique is to overcome the ethical problems when designing experiments in education. If a technique is trialled on two groups of students and the difference between their attainment is measured then one group will be advantaged/disadvantaged, which is not good practice.

Assuming that a cohort of students find two courses:-

* to be equally difficult
* are assessed to an equal standard
* are taught to the same standard
* have an equal access to learning materials.

then it would expected that the students would get exactly the same results in both courses, i.e. the results from these two courses would form a straight line with *a*=0 and *b*=1.

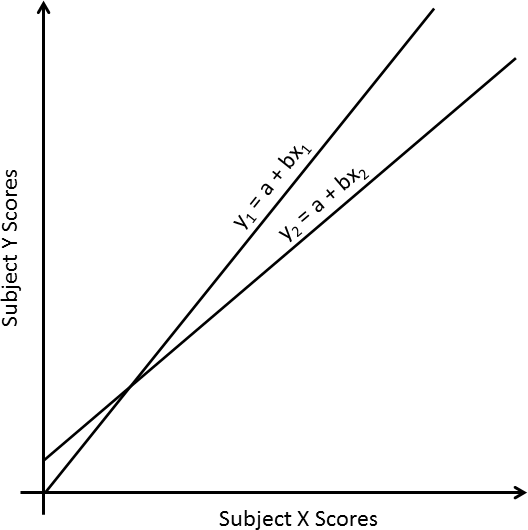


The actual regression line will be different - and that needs assessing before the introduction of any new materials or methods. The data for as many previous years as possible should be examined across as many courses as possible. A data group could be formed for each course combination of those students who completed both. The correlation and regression line could then be calculated for each course pairing for each academic year.

Changes in teaching method or changes in teaching staff would show up as changes to the slope, *b*, between years since student preference for the two courses could be altered. However, a complicating factor would be that the introduction of new materials or teaching methods could also alter student preferences. A longitudinal study would be expected to show up each of these factors as changes between year groups which are static and those which are dynamic are examined.

Student perception aside, the introduction of new materials would be expected to keep the slope of the line the same (the brighter and more dedicated will continue to achieve better than those less able and less motivated). Therefore the intercept could be expected to show the effect.

If the changes to *a* are statistically significant then the new materials would have had an effect. The degree of that effect (or not) could be shown by its distance from the expected value of *a*.



Another measure could be the effect on the correlation coefficient. In an ideal world, it would be found that *r* = 1. For that to happen all the students would have to achieve the same normalised score (taking the regression line into account) for each course – and that would be suspicious since students vary.

These differences in performance for students between units (learning dissonance) should diminish if the new materials or teaching methods an effect. It will not affect those below the regression line as much since these students are performing better in comparison course. However, those who have more difficulty with the altered course will have access to new materials and/or alternative learning modes and would therefore be expected to improve. The degree of improvement in r2 × 100 (or not) would be another test of the effectiveness of changes in the teaching environment.

Finally, the degree of learning dissonance between courses:-

would be expected to reduce as standard of teaching becomes more uniform. Statistically large discrepancies (a term that would need to be quantified by this study) might be an area of concern.

**Practice: The 2010/11 academic year results**

The tables in the next section show:-

**Table 1** : The number of students who took both the course. The numbers are not the same for each cell since students from other programmes may also have participated.

**Table 2** : The Pearson Product Moment Correlation Coefficient (r) between the two courses for the scores in each of these courses. Any scores starting with L were deemed to be 40.

**Tables 3 and 4** : The intercept and slope for the linear regression line between these two courses.

**Table 5** : The theory outlined in the previous section hypothesises that the relationship between these scores would be linear. This has been tested by sorting the scores for the y- and x-variables and then using linear relationship to estimate the y-variable. The Durbin-Watson d-statistic should be near to 2 if the relationship is linear. Ten of the results (indicated with \*) show a significant level of autocorrelation which indicates a non-linear relationship. This has not been explored further yet.

The row of interest is the top one for COMP1148, since that is the course on which the proposed system will eventually be trialled.

If the new system were to have an effect then the correlation coefficient would be expected to improve because the results would bunch more at one end. If both the mean score () and the correlation coefficient (*r*) increase then the effect would have been beneficial.

Mutual correlation could be a problem for this approach. If COMP1148 becomes more understandable then students may feel generally more motivated and achieve better scores in other units too. The degree of effect can be estimated using the changes in *r* relative to COMP1148.

The only change to COMP1148 this academic year was the replacement of Kate Finney. The materials were the same and the method of marking was the same. It will therefore be an interesting exercise to see whether this one change made any difference that detectable statistically. This can be tested once first year scores have been finalised in September 2013. This test will show whether the statistical tool described here has any potential.

**Potential further interesting areas**

1. Could this approach be used to assess the effectiveness of new teaching materials as well as approaches to teaching?
2. Is there anything which can be learned about the level and effectiveness of other courses? Could it, for example, highlight areas of excellence or areas for improvement?
3. Could this approach be used in situations where non-numeric results are achieved, for example, fail, pass, merit and distinction as on National Diplomas? If so, one would have to use Spearman’s Rank Correlation Coefficient instead.

**Cross-Course Statistics for 2010/11**

**for those taking courses on the year 1 of the Computing programme**

**Table 1**: n (Sample size - the number taking both courses)

COMP1152 COMP1587 COMP1588 COMP1589 MATH1110 MATH1111

COMP1148 260 337 152 189 193 169

COMP1152 250 141 107 176 152

COMP1587 148 183 189 164

COMP1588 0 136 137

COMP1589 55 30

MATH1110 164

**Table 2**: r (Pearson's Correlation Coefficient

COMP1152 COMP1587 COMP1588 COMP1589 MATH1110 MATH1111

COMP1148 0.8484 0.5946 0.7911 0.6673 0.6286 0.8004

COMP1152 0.6243 0.8252 0.6835 0.6578 0.8093

COMP1587 0.7046 0.3966 0.6741 0.6203

COMP1588 \*\*\*\*\*\* 0.7156 0.8502

COMP1589 0.3333 0.4435

MATH1110 0.8146

**Table 3**: a (The intercept)

COMP1152 COMP1587 COMP1588 COMP1589 MATH1110 MATH1111

COMP1148 16.014 36.295 10.460 30.652 31.353 10.057

COMP1152 29.412 0.090 24.484 20.785 -1.600

COMP1587 -2.544 42.100 19.897 7.610

COMP1588 \*\*\*\*\*\* 27.796 5.754

COMP1589 35.957 32.595

MATH1110 -5.211

**Table 4**: b (The slope)

COMP1152 COMP1587 COMP1588 COMP1589 MATH1110 MATH1111

COMP1148 0.685 0.514 0.713 0.610 0.496 0.770

COMP1152 0.681 0.914 0.780 0.686 0.980

COMP1587 0.816 0.387 0.614 0.715

COMP1588 \*\*\*\*\* 0.619 0.908

COMP1589 0.392 0.524

MATH1110 1.000

**Table 5**: d (Durbin-Watson d-Statistic)

COMP1152 COMP1587 COMP1588 COMP1589 MATH1110 MATH1111

COMP1148 1.4773 1.3615 2.0456 1.5980 1.5154 1.8166

COMP1152 1.7927 2.0437 2.1814 1.9259 2.0762

COMP1587 1.7225 1.5370 1.5538 1.6947

COMP1588 \*\*\*\*\*\* 1.2925 1.5445

COMP1589 2.0089 2.0001

MATH1110 2.2303

**Table 6**: ld (Learning dissonance)

COMP1152 COMP1587 MATH1110 MATH1111 COMP1588 COMP1589

COMP1148 6.060 8.287 8.422 7.011 8.518 7.671

COMP1152 7.965 6.668 6.670 9.179 7.864

COMP1587 9.770 9.538 7.367 9.913

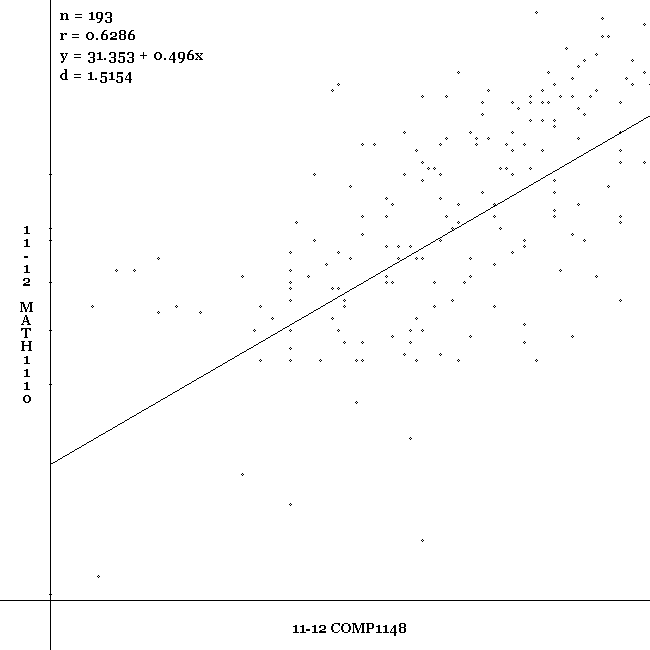
MATH1110 \*\*\*\*\* 7.957 7.110

MATH1111 8.556 8.014

COMP1588 6.190

Note: Some combinations (marked \*\*\*\*\*) are mutually exclusive.

**A Sample Graph**



**Creating the Data Files**

There are several problems with the files from BannerWeb.

1. The names and enrolment numbers of the students are shown.
2. Each student record runs over four lines of text.
3. Each record contains information that is not relevant.
4. The full program title appears multiple times.
5. Students on programs unrelated to this study appear, e.g. collaborations.

Problems 1 to 4 can be overcome by creating program which outputs five text files; a student file, a course file, a course tutor file, a programme file and a scores file. Problem 5 will be addressed in the data analysis program by excluding them from the results.

Each student will have an ID (which is an MD5 encryption of their enrolment number) and student name. The student ID will become the new version of the enrolment number and name of the student when the data is released.

* studentID (the MD5 encryption of their enrolment number)
* enrolmentNumber
* fullName

Each course will have an ID (a random number), the course code and the level. Again, the course ID will replace the course details when the data is released.

* courseID (a unique random number)
* courseCode
* level

Each course tutor will have an ID (a random number), the course code and the level. Again, the course tutor ID will replace the course tutor details when the data is released.

* courseTutID (a unique random number)
* courseTutName

The programme file will have an ID (a random number), the programme code and the programme name. Again, the row ID will replace the programme details when the data is released.

* programID (a unique random number)
* programmeCode
* programmeName

Finally, the results file will contain all the information to be used in any analysis. This contains the only data that would be released.

* scoreID (a unique random number)
* studentID (the MD5 encryption of their enrolment number)
* courseID (a unique random number)
* courseTutID (a unique random number)
* programID (a unique random number)
* academicYear
* semester
* level
* noOfCredits
* score

This final file would be the only file to be released into the wild. Having this as a flat file of text makes the data format independent. The layout of the file would allow the data to be analysed in many ways whilst protecting the anonymity of the student, the lecturers, the programmes and the courses.

### Appendix B - The Layout of the Data

The original data obtained from BannerWeb has the following format:-



This will be converted into the following format as text files (for ease of programming):-

