Predictors of first year computing science student failure

Maxi project March 2003

Sarah Rebecca Black

ABSTRACT

There is an abundance of research on Higher and Further education student attrition issues. This study addresses student attrition in the Computing Science course at the University of Glasgow. Following up a project by Roddan (2002), a number of variables were investigated with the aim of building a predictive model of students at risk of failure. Instruments utilising the Tinto Student Integration Model (1975), indicated that academic integration factors explained a significant amount of the variance in first year student exam performance. Results are presented and discussed, and recommendations for further research are made.

N.B. many of the appendices are missing from this version, although the questionnaire is included.

Introduction

Student failure and attrition from university courses is a significant concern. Recent statistics printed in The Guardian newspaper (2001) report that more than one in six undergraduates in the UK dropout of courses early. There is a considerable human cost and the cost to the taxpayer is £200million drawing the attention of MPs. This paper examines student failure and attrition in the first year of a large undergraduate computer science course at the University of Glasgow in the academic year 2002-2003. In 2000 42.2% of the first year students failed the programming module. The department are concerned about this and are keen to discover why so many fail. The aim of this study is to identify factors which could provide an early indication of students at risk of failing, so that it might be possible to help and prevent this. This study primarily follows up from a study by Matt Roddan (2002). Before the present project is explained any further it is important to review the research up to this point.

Background Theory

The major and most comprehensive, conceptual model of student drop-out is Tinto's Student Integration Model (1975). Tinto's theory has been, and continues to be the foundation of much of the research carried out on student departure. In 1975 Tinto described his theory of college student departure as follows: Students' individual characteristics play a role in the departure process. These individual characteristics include, individual attributes (for example, ability, race and gender), family background characteristics (for example, parental educational level, parental income), and pre-college/university schooling experiences/ academic background, (e.g., records of achievement, grades achieved). These characteristics affect the level of initial commitment to the institution and to the goal of graduation, these commitments affect the student's degree of integration into the academic and social systems of the college or university. The greater the student's level of academic and social integration the greater the likelihood the individual will persist in college. (Bray, Braxton, Sullivan 1999).

Students' social and academic integration-investigating Tinto's model (1975)

Tinto's Student Integration Model (1975), described earlier, is the single most referenced theory in the research of student drop-out. Tinto's model is often investigated in various higher and further educational institutions, frequently providing significant results. Elkins, et al (2000) examined Tinto's concept of separation on first year university student persistence decisions. Tinto describes, "the first stage of the college career, namely separation, requires students to disassociate themselves, in varying degrees, from membership in past communities" Tinto (1988). Data analysis revealed interesting results which provided support for Tinto's concept of separation influencing student persistence in this sample.

A longitudinal study by Thomas (2000) looks into student social integration and persistence. The results showed that student acquaintances had positive effects on grade performance and persistence.

When examining the literature much research focuses on the impact of stress on student departure decisions. Bray, et al (1999) study student departure focusing on how students deal with stress and its impacts on social integration. The results indicated the greater the students use of positive coping strategies, the greater their degree of social integration into their university. Bray et al advised that institutions should include stress management into classroom learning.

Torres and Solberg (2001) evaluated a model of college outcomes, which included four constructs: academic self-efficacy, stress, family support, and social integration. Results indicated that self-efficacy directly predicted social integration, persistence intentions and stress.

Montmarquette et al, (2001) also attempted to find predictors of persistence from longitudinal data on student enrolments. Following data analysis the variables which explained early drop-outs were related to a non-traditional class size effect (a group of more than 87 students). Smaller classes increased the probability of persistence over an early drop-out.

The current project intends to gather data on student perceptions of informal and formal contact with the staff of the department, Mannan, 2001, used the Tinto model to make an assessment of the academic integration with staff, and social integration as perceived by students of the University of Papua New Guinea. Mannan used a research instrument based on the Tinto (1975) model. Significant differences were found between the current and ideal perceptions of student groups in respect of a) students' informal contact with faculty and b) faculty concern for student's development and teaching. The methods of Mannan's study could be used as guidelines for assessment in other institutions.

Tinto's theory has yet to be applied to this course. The computing science course at Glasgow University has a large first year class. For this reason it is possible that integration may be an issue. Students may find it difficult to feel a part of such a big class and to integrate socially with other students despite smaller tutorial groups. In addition, previous first year students' expressed feelings of a lack of academic integration into the course, feeling that staff were uninterested in their progress, and it was not made clear to students, in the first few weeks of term, about how to get help if needed. The class is largely male dominated, staff in the department suggest that girls more often ask questions in tutorials and that boys are less likely to do this, so do less well at the course.

Using questionnaire items utilising Tinto's variables, this study will discover whether its concepts can be usefully measured in practice. As Tinto's model is designed to be a longitudinal one, the questionnaire items will be asked on two separate occasions, to determine whether there are any changes over time. It is clear from the studies above that it is not only Tinto's variables which have an influence on student departure. For this reason the current study measures Tinto variables and a wide range of other variables which may influence student departure in this sample. Although Tinto's theory is the major one, many believe that too much research on student dropout focuses on this model which is quantitative in nature, Yorke criticises, "Tinto's model makes assumptions about how students reach dropout decisions, without consulting any students as to whether these assumptions hold true" (2001). The following two studies are quantitative and consult students in order to gain an insight into dropout.

Gracia and Jenkins (2002) employed qualitative research methods, these comprised of a series of in-depth semi- structured interviews with a sample of 42 students. Their aim was to understand through student experience why some students pass and others fail. The results identified differences between students who failed and those who passed. The characteristics common among failed students included; "Negative focus of reasoning and the impact of affect". These students provided 'negative' reasons to support their initial study choices. Another characteristic among these students was, "Patterns of participation". Interviews revealed that failed students described more reticence to participate than students who passed, they described feelings of lack of confidence to do so. This study reveals interesting results about the cognitive aspects of failure in students and highlights the importance of qualitative research.

Much research literature focuses on aspects of the teaching and learning students encounter at university. Braxton, et al, (2000) investigated whether forms of active learning, such as group work and class discussions, influence social integration, subsequent institutional commitment and departure decisions. In their sample, class discussions positively influenced social integration which, positively influenced institutional commitment and persistence. Braxton et al. conclude that the role of active learning positively influences student persistence.

Karen Hinett (1998) carried out research across 26 disciplines. Following data analysis, Hinett suggests the role of dialogue helps students perceive some ownership and control over their learning. The feedback offered to students was important to student learning and positively influenced their commitment to the institution. This may have a particular application to first year students who are unaware of the standard they should be achieving.

Literature on the attrition of Computing Science Students

There is an abundance of literature on student dropout in general, but much less on the dropout of students from computing courses and specifically the failure of these students to succeed at the computer programming aspect of such courses, which is the interest of this study.

Much of the literature attempts to improve the teaching of programming and discover better ways to make sure the students grasp it. The criteria for course entry has also been scrutinised, although it seems that this is only weakly linked to student failure in the subject.

Tony and Jenkins (2001) suggest that the previous experience students have before they begin the course is important in terms of success or failure. However, often the literature concludes that previous computing experience does not influence student success and persistence, but as Jenkins explains, "even those students...with an educational background in computing may well not have programmed to any meaningful extent, as programming is dropped from some 'A'-level courses". The current study will collect data on the students' previous *programming* experience, not just computing experience in general, to address this issue.

Stephens and Creaser, (2002) carried out a longitudinal study of student computing experience over five years. The results describe the actual trends in prior experience of new information science undergraduates in Information Technology. Their questionnaire asked students to rate their experience, ability, knowledge and confidence on a variety of computing topics. Stephens and Creaser conclude that although student experience of general software is increasing, prior experience with programming is not. The students rated themselves as having less ability at programming than with other computing skills. The authors recommend the regular testing of in-coming students to assess their real computer needs. Hagan and Markman (2000) agree suggesting that students with experience in at least one programming language perform significantly better in assessments.

A project by Matt Roddan (2002) focused on the Level 1 Computing Science students at Glasgow University, addressing the department's high rates of level 1 student failure on the course. Roddan's aim was to measure variables and to correlate these with students' exam performance, to identify factors which might provide an early indication of a student who is at risk of failing. Roddan (2002) attempted to find a correlation between students' entry point scores to the university and their exam results. This variable did not account for a significant amount of the variance in exam score. However, it is possible that entry point score was not fine-grained enough to gain important information about the predictive power of prior experience.
Therefore with regards to the current project it was decided to decompose entry point score into more specific terms. Having interviewed second year students of this course, it was clear that previous experience of programming alone does not predict success. Many of the students who were interviewed had no previous programming experience, some of whom achieved a grade A at the end of first year. It is possible that the benefits of previous programming experience are simply to do with thinking in programming terms. Previous students suggested that learning to program requires a new way of thinking, success requires students to experiment with it and get lots of practice. Perhaps these are the factors which make prior programming experience valuable.

From Roddan's investigation, the only factor with good predictive power was students own self-estimate of how well they understood the course material, this accounted for 49% of the variance in exam performance. It would be beneficial to identify when students can accurately judge their own understanding as Matt discusses, "it may well be the case that the students themselves have a good instinct that they are at the risk of failing long before they can be identified by staff." There are however, some methodological concerns which reduce the significance of this result. The students were asked to rate their level of understanding after they had sat the January exams and some students had already received their results. This will have biased the results. It is of the interest of this study to examine this factor in detail. The idea is to identify students at risk of failing the course as early as possible so that help can be provided. This study may discover that Roddan's findings were due to temporal variables, in which case the students' self-estimates would not explain a significant amount of the variance. However, when a significant variable is discovered it is important to try and replicate these findings.

Students will be asked throughout the first term how well they understand the course material. Previous students of this first year course often recall that it is important to get to grips with programming in the first few weeks, to be able to apply this knowledge later on when the course gets more difficult. Perhaps students which report low levels of understanding in the first few weeks of the course will perform more poorly in the exams as they haven't had the basic skills to understand the more difficult material later on. The aim is to identify when students' self-assessments become predictable and correlate with exam performance. If this study yields such information, it would provide a way of identifying students at risk, who are not understanding the course material, and who could then be offered some extra support by the department.

Roddan also collected information on the amount of revision each student did for the January exams. This measure did not account for a significant amount of the variance in exam score. Roddan suggests that this result is due to the particular learning processes needed to acquire computer programming skills. Course leaders in the computing science department also take this view as does Jenkins (2001). Jenkins states, "Programming is certainly a complicated skill to master, and learning to program is correspondingly complex". They believe that it is not possible to simply revise computer programming for an exam without having understood it from the beginning of the course. Roddan's result would suggest that it is understanding of programming, not the amount of revision done which is vital to passing the programming exams.
With this in mind this study will ask students how much independent study they do for this subject, this will be in the form of items on a questionnaire. It will be of interest to discover whether increased study of programming will lead to better exam results within students.

Alongside the variables identified by Roddan's project, this study investigates variables which, from the literature and departmental contributions, were considered as possibilities to explain a proportion of the variance.

From talking to staff in the computing science department, it was suggested that level 1 tutors may themselves, be able to make a good estimate of how well their students will do in the January exams. Therefore, the Level 1 tutors were asked to predict a grade for each of their students. If tutors' predicted grades correlate with students' exam results, it would be possible for tutors to identify failing students and to offer them the extra support they may need.

It is hoped that by measuring the variables described in the sections above, those which are instrumental within student failure, can be identified to build a predictive model of failure appropriate to Level 1 Computing Science students at the University of Glasgow.

An Overview of the Level 1 Computing Science Course at the University of Glasgow

There is a total of 324 students in the level 1 class in the academic year 2002-2003. The course is broken up into two modules, CS1Q and CS1P, each worth 20 credits, (students must attain 120 credits in their first year by taking three subjects in total). The CS1P module teaches the principles of programming, and students learn to program in Ada95, this is the module from which the data for this study will be collected, as students tend to have most difficulty with this section of the course.

Two class tests and lab exams, sat in January and June make up the course assessment. The class test is a written examination, and assesses student's understanding of programming. In the lab exam, students are given a program to write a few weeks beforehand, which they must reproduce under exam conditions on a computer. Students must attain a grade G overall to complete the module and gain their credits, and a grade C to progress onto Level 2 of the computing science course. Unfortunately, the scope of this project did not allow for collection at the June exam results to determine which students have failed, or for collecting data on which students drop-out. Therefore it concentrates on the results from the exams sat in January 2003.

Method

Permission for this study was obtained from the Department of Computing Science at the University of Glasgow. The Psychology department granted ethics approval.

Participants

The participants were the full-time, first-year Computing Science students at the University of Glasgow in the academic year 2002-2003. The students who volunteered were briefed about the study and advised of the basic aims of the experiment. They were assured that all of the information they submit would be treated in the strictest confidence.

Instruments and Procedure

The four survey instruments were carefully chosen and designed to gather quantitative information pertinent to retention issues and causes of student departure or failure. Permission was granted by the CS1P course leaders to administer each of the instruments described.

1. Understanding Question- Students' self-estimates of their Understanding

This instrument was designed to follow up the finding of Roddan's earlier study. The aim is to replicate Roddan's finding that students' self-estimates of their understanding predicted 49% of the variance in exam results. When do students estimates become predictive of exam result? This could provide a way of identifying students at risk of failing the CS1P course.

Students were asked, " How well do you feel you understand the material on this course (CS1P) at this time?" They indicated their level of understanding on a scale from 1 ("I have no understanding") to 9 ("I understand everything"). During the third week of the first term a lecturer asked this question in one of the students' CS1P lectures. The question was asked regularly to see whether students' levels of understanding changed, and when such a change occurred. The students were asked the question at the beginning of the lecture and were asked to indicate their response by pressing the appropriate key on their personal handsets. This data was collected automatically by receivers connected to the lecturer's computer and was then presented to the investigator in a Microsoft Excel spreadsheet package.

2. Prior Experience and Integration Questionnaire Appendix 1

This questionnaire consisted of 78 items. The first ten items ask the students for their matriculation number, age, intended degree, and prior computer and programming experience. Levels of academic and social integration of the students were assessed by the next 42 Tinto items on the questionnaire, described previously. Finally, the last 3 items of the questionnaire were open questions chosen to gather qualitative information on the students' feelings towards different aspects of the course.

The questionnaire was piloted on a small sample of people to ensure that the Level 1 students would understand and complete it without any problems. It was distributed to the students during the second week of the first term via the CS1P tutors who were briefed on the purpose of the study. They were asked to handout the questionnaire in the last ten minutes of their tutorial. The tutors then collected the questionnaires once the students had completed them and left them in the Computing Science department front office where the investigator would collect them. 218 completed questionnaires were obtained.

3. Study Habits and Integration Questionnaire Appendix 2

There are a total of 72 items in this questionnaire. The first 7 items ask students for their matriculation number, age, and information on their study habits. The rest of the items on this questionnaire included the Tinto Integration items identical to those measured by the Prior Experience and Integration Questionnaire. In week 5 of the second term the students completed this questionnaire. Having been granted permission by the Course leaders, the project investigator distributed and collected the second questionnaire during the last ten minutes of the students' tutorials. 174 completed questionnaires were obtained from this data collection.

Both of the questionnaires above measured the levels of social and academic integration of the students. 42 exploratory items on these questionnaires were designed to utilise the Tinto Student Integration Model (1975). 16 items measured academic integration and 26 items were designed to measure the social integration of the students. The statements asked about the students' experiences and perceptions of the CS1P course, to which respondents indicate their degrees of agreement or disagreement using a 7-point scale. The students could also tick the non-applicable box for any item. Also, included were two questions asking the students to give the number of students they know, and the number of personal interactions they have had with a member of departmental staff. It was decided to measure these items twice, as it is likely that levels of integration change over time, affecting the potential predictability of this variable. In addition, Tinto's model is designed as a longitudinal one, so it was appropriate to measure variables utilising it over a length of time. The aim of these instruments was to discover whether the CS1P students' levels of academic and social integration would be predictive of exam results.

4. Level 1 Tutor Predictions Appendix 3

During the initial design of this study, departmental staff of the CS1P course had suggested that the Level 1 tutors may be able to predict which of their students would pass or fail the January exams. This instrument was designed to test this theory. If tutors could predict students at risk of failing the course, they could identify such students and offer them extra support. In the last week of the first term each of the 22 CS1P tutors were given a sheet, on which it had the names and matriculation number of the students in their tutorials (Appendix 3). CS1P tutors were given a sheet to fill in and asked to predict a grade from A to N for each of their students. Following the Christmas break, the tutors were asked to give their completed sheets to one of the CS1P course leaders from whom the investigator collected them. Predictions from eight tutor groups, for 84 of the Level 1 students were obtained.

5. Departmental Data

The Department of Computing Science holds information about every Level 1 student. The department provided the investigators with the January CS1P lab exam and class test results (the dependent variable), unless a student had not allowed their result to be known. Finally, during week 17 of the second term the results of the January CS1P exams the students sat in week 14 of the same term, where gained from the department's administrative staff.

Results

The following section will describe the most pertinent results gained from the data collection.

<u>CS1P January Exam Results - Descriptive Statistics</u>

Table 1

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
LABEXAM	281	20.00	.00	20.00	13.7651	5.5954
CLASTEST	285	97.00	3.00	100.00	60.2175	19.4520
Valid N (listwise)	278					

The Lab Exam was out of 20, the Class test was out of 100. The mean score in the Lab exam for each student was 14 out of 20, or approximately 68%. The mean score in the Class test was 60%. The students on average get a higher score on the Lab Exam than on the Class Test. This supports the general view held by lecturers that the class test is a more reliable test of understanding, as it is easier for poorer students to get a reasonably high mark in the lab exam.

<u>1. Understanding Question Data Analysis</u>
In order to follow up a lead from Roddan's project this study collected data on student's self-estimates of their understanding. The purpose of this data collection was to discover whether student's self-estimates are predictive of exam result, and if so, when they become predictable. This data was collected in one of the CS1P lectures.

Self-estimates were collected from 262 students, however not all of the students responded at each session. See the table in Appendix 5 where the number of data obtained and the number of missing data can be found for each session.

There were a total of 13 sessions in which the Understanding Question was asked. The dates of the sessions can be found in Appendix 4. There is a lot of missing data from the sessions. However it is still possible to analyze the data collected, in terms of the results of the class test. The following analyses were carried out.

Descriptive Statistics

The minimum and maximum rating of understanding for each student, and for each session was calculated. The session with the lowest mean rating of understanding for all students is session 9 = 6.347. This rating is quite high and suggests that students understood the material well. Session 19 had the highest mean rating of understanding for all students = 7.667. There is not a large range between the minimum and maximum ratings. The descriptive statistics for each session can be seen in Appendix 6. These results must be interpreted with caution as there is a lot of missing data so these results may not be representative of the whole class. Also there was often a large range between minimum and maximum rating within each session, (1-9). For each student as well as the minimum and maximum being calculated, the session in which the minimum occurred was gained and the maximum rating after this minimum were obtained also. Correlations between these variables and class test were not significant and can be found in Appendix 7. The mean place of minimum rating by each student, occurred in session 9.

Correlations between session data and class test results

Having gained these descriptive statistics of the data the next step in the analysis was to look for correlations between this data and scores in the class test. The class test was used as the dependent variable, and not lab exam, because the class test is deemed a more reliable measure of understanding by CS1P staff. Due to the large amount of missing data and the random distribution of the data, individual correlations between each session and exam result were calculated to achieve the most information from this data collection. Table 2 below shows the results of the Pearson's correlation coefficient test, a perfect correlation would have a value of 1. A significant correlation produces a p-value of <0.05.

Table 2

Session Number	Pearson's Correlation Coefficient	Significance (0.05 level)
5	0.193	0.266
6	0.200	0.007
7	0.134	0.100
8	0.360	0.040
9	0.376	0.000
10	0.321	0.096
11	0.432	0.000
12	0.565	0.001
14	0.375	0.000
18	0.404	0.000
19	0.302	0.239
22	0.340	0.000
23	0.586	0.008

The table of correlation coefficients above shows that the sessions nearest the beginning of the term sessions 5-7 have the lowest correlations with class test result. The highest significant correlation between understanding rating and class test result occurs in Session 23 correlation = 0.586, p-value = 0.008. The student's ratings correlate highest with their class test result in the final session in the last week of the first term. The second highest correlation occurs in Session 12 =0.565. The lowest correlations are in the beginning sessions 5-7, the highest correlations occur in the later sessions, 12 onwards. This suggests that student's self-estimates become increasingly correlated with class test result over time. The next step in the analysis of this data would have been to add it into a regression equation with other significant variables, unfortunately, due to the sparseness of this data, such analyses could not be done.

2. Prior Experience Data Analysis

This data was collected by distribution and collection of the "Prior Experience and Integration Questionnaire" (see Appendix 1) in the first term. The purpose of this instrument was to gather fine-grained information about student's prior computing and programming experience. It is of interest to discover whether prior experience of this kind has a significant effect on the Level 1 student's exam results. Prior experience data was collected from 125 students. For the descriptive statistics see Appendix 8.

MANOVA (Multivariate Analysis of Variance)

A MANOVA was calculated for each of the prior experience items on the questionnaire (see Key below).

Key: CSQUAL = Computing Science Qualification; any previous CS qualification 1= Yes 0=No.
HRSPROG= Number of hours of programming experience
GCSKILLS= General computer skills rating from 1(novice) –7(expert)
PROUSED= Used programming languages before 1=Yes 0=No

The MANOVA statistic was chosen as the appropriate one to test this data. The MANOVA looks at how the independent variables (above) interact with each other, and what effects these interactions have on the dependent variables, which in this case are the results of the CS1P Lab exam (%) and Class test. The results of the MANOVA return p-values, a p-value below 0.05 means that there is a significant effect. According to Pillai's Trace criterion, none of the prior experience variables had significant p-values below 0.05.

CSQUAL, ($F(2, 109) = 0.9163$, $p= 0.404$), HRSPROG, ($F(6, 220) = 0.645$, $p= 0.694$), GCSKILLS ($F(10, 220) = 0.852$, $p= 0.579$), PROUSED, ($F(2, 109) = 1.859$, $p = 0.161$). For the full MANOVA results see Appendix 9. Statistically these p-values indicate none of these variables had an effect on class test or lab exam, therefore no further analyses were carried out. Not all of the variables above had the same missing variables, to ensure the MANOVA was valid, separate ANOVAs were ran, which also produced the same insignificant results.

3. Study Habits Data Analysis

Data was collected on student's study habits using 5 items on the "Study Habits and Integration Questionnaire" (see Appendix 2) which was administered in the second term. 125 cases of student data was obtained from the distribution of this questionnaire. The purpose of this instrument was to discover whether the particular study habits of a student have a significant effect on their January CS1P results. Appendix 10 illustrates the descriptive statistics for this data.

MANOVA (Multivariate Analysis of Variance)

The MANOVA statistic was calculated for each of the 5 study habit items on the questionnaire.
Key: INDWK= whether the students do any independent study for CS1P 1=Yes 0=No
INDHRS= number of hours of independent study in last seven days
WHERE= where the students normally study 1= university campus, 0= home 2= both
TOGETHER = how often the students study together with other students
FACILITI= whether the students have the facilities they need to study at their term time address

The MANOVA returned p-values of the significance of the effects of each of these variables. Using the Pillai's Trace criterion. These are as follows; INDWK, ($F(2, 95) = 1.158$, $p= 0.318$), INDHRS, ($F(28, 192) = 1.229$, $p= 0.173$) WHERE, ($F(4, 192) = 0.536$, $p= 0.709$), TOGETHER, ($F(10, 192) = 1.213$, $p= 0.285$), FACILTI ($F(4, 192) = 0.335$, $p=0.854$). For the full MANOVA results see Appendix 11. All of the p-values for these items are greater than 0.05, which means they are not significant. These results indicate that none of the study habit variables have a significant effect on the CS1P lab exam or class test in January. The descriptive statistics (Appendix 10) add further support for the insignificance of these results with the range between the variables often being very small.

4. Integration Data Analyses
42 items on both the "Prior Experience and Integration Questionnaire" and the "Study Habits and Integration Questionnaire" were designed to collect data on the CS1P student's levels of academic and social integration. During the analysis of this data, it was decided to use the integration data collected in the second term and not to analyse the data collected in the first term, this decision was made as following factor analysis they explained very little of the variance, this was perhaps because the questionnaire was distributed early in the first term and some of the items on the questionnaire were not applicable. The integration items from the second collection in the second term were used. The analysis of these results is as follows. The integration items consisted of 16 items measuring academic integration and 26 items measuring social integration.

Analysis of Academic Integration Items
A principal components factor analysis was carried out on the 16 academic integration items (see Appendix 12). This analysis attempts to establish which linear components exist within the data and how a particular variable might contribute to that component. In other words this analysis will group academic integration items together, which measure similar components. The rotation component matrix revealed the items, which were statistically grouped together into 3 factors. Although the items on the questionnaire were grouped together statistically, this does not take into account what they are conceptually. To see whether these questionnaire items are actually similar in real-life, need to take into account what these item statements are in the questionnaire. The first factor to be identified named 'Learning' is made up of the following items:
Q2TA14 - "So far I am understanding the material as well as I want to".
Q2TA1 - "I am getting on very well with my studies in CS1P"
Q2TA15- "I am getting good enough marks"
Q2TA13- "I feel comfortable with the amount I am learning in this course so far."

These items go together conceptually, they are concerned with student's feelings towards their learning. Therefore this supports the rotation and principal component analysis.
The second factor identified and named 'Course Attitude -positive' follows:
Q2TA9- "I find CS1P interesting"
Q2TA6- "Getting a good grade in this course is important to me".
Q2TA12- "I am enjoying studying CS1P"
Q2TA8- "I like the learning activities I am asked to complete in this course"
Q2TA4- "Studying CS1P is useful"

These items identified as a factor by the principal component analysis also go together conceptually, they are concerned with a positive attitude towards the academic aspects of the course.

Finally, the third academic integration factor consists of the following items and is called 'Course Attitude - negative':
Q2TA5- "This course prevents me from engaging in activities that I like"
Q2TA2- "This course demands things of me that I don't like"
Q2TA10- "This course blocks me from doing things which are important to my learning"
Q2TA17- "Apart from getting a qualification there is no value in coming to this university".

This items are to do with a negative attitude towards the academic aspects of the CS1P course, they go together conceptually as well as having been put into the same factor by the principal components analysis.

To ensure that these factors were reliable, that there is an internal consistency within these factors a Reliability Analysis was carried out on each factor (see Appendix 13). The Reliability Analysis allows you to study the properties of measurement scales (our factors) and the items that make them up. The model used in this analysis was the Alpha Cronbach Model. This is a model of internal consistency based on calculating the inter-item correlation coefficient. If the Alpha value is 1 then items within a factor are perfectly correlated, and are consistent, making the factor reliable.

Results of Reliability Analysis- Academic Integration Factors
'Learning' - Alpha= .8610 '
Course Attitude –positive' - Alpha = .7879
'Course Attitude- negative' - Alpha = .4418
The factors 'Learning' and 'Course Attitude-positive' have high Alpha coefficients and are therefore reliable. However the third academic integration factor 'Course Attitude-negative' has a low Alpha coefficient, and is therefore not reliable.

Having carried out the principal components analysis and the reliability analysis, three factors, which explain 50% of the variance in the academic integration item scores have been identified. These factors will be put into a regression with the factors identified from the social integration item data to build a predictive model of exam failure.

Analysis of the Social Integration Items
The same analyses of the academic items are performed on the social integration items to discover the factors which explain the most variance within these scores. A principal components analysis was carried out on the 26 social integration items on the questionnaire (see Appendix 14). As there are more social items than academic ones, and often items which overlap in meaning, this factor analysis was less concise.
5 factors explained 56.6% of the variance.

The first social integration factor named 'Experience' comprises:
Q2TS40 – "Going to university and being a student helps me to get on with other people…"
Q2TS41- "Going to university makes me fit in better in life outside the university."
Q2TS33 – "Going to university fits in with the kind of person I want to be"
Q2TS34- "I feel comfortable being a student in the UK today"
Q2TS37- "Being at the University of Glasgow is impressive to others"
These items are mostly about the experience and perceptions of university students have. To ensure that this factor is reliable a Reliability Analysis was performed. =.7742 (see Appendix 15).

The second social factor named 'Friends' is made up of the following:
Q2TS27- "I have made as many friends as I want at university"
Q2TS38- "I fit in with other students in the university"
Q2TS22- "I like the conversations I find myself having with students at university"
Q2TS26- "I enjoy the social activities that other students suggest"
Q2TS28 – "I feel comfortable around campus, the departments, in lectures etc."

The majority of the items in this factor do relate, they are concerned with students relationships with other students. To ensure that this factor is reliable a Reliability Analysis will be performed (Appendix 15). The initial Alpha value returned by this analysis, alpha = .7002 so this factor has a high reliability.

Lastly the third social factor includes the items below, as is named 'Staff/Department':
Q2TS16 –" I feel comfortable being a student at this University"
Q2TS20 – "Getting to know staff and students is useful to me"
Q2TS30 – "I feel comfortable approaching staff whenever I feel I need to."

These items are about students social integration with the staff and department. To ensure reliability the Alpha value = .676. (Appendix 15). The two other factors identified by this principal component analysis did not elicit significant reliability values and therefore will not be included in any further analyses.

Regression Analysis of Integration Items
Having identified three academic integration factors ("Learning", "Course Attitude +ve", and "Course Attitude -ve) and three social integration factors, ("Experience", "Friends" and "Staff/Department"), the next step in the analysis of this data is to discover whether scores on these factors are predictive of student's class test result. (Class test used as it is the best measure of student understanding.)

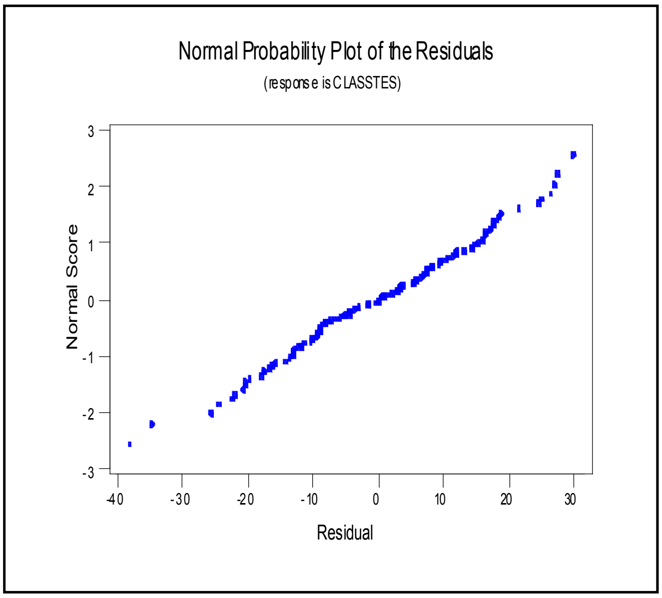
In order to build a predictive model of exam result the academic and social factors are put into a regression with class test result. The result of the Best Subsets Regression of the 6 integration variables follows.

Best Subsets Regression: Class test vs. 'course attitude –positive', 'course attitude –negative', 'Learning', 'Experience' , 'Friends', & 'Staff/Department'

Vars	R-Sq	R-Sq(adj)	C-p	S	A	B	C	D	E	F
1	36.5	36.0	2.2	14.138			X			
1	7.4	6.6	57.3	17.071		X				
2	37.8	36.8	1.7	14.050	X		X			
2	37.3	36.2	2.7	14.107			X		X	
3	38.8	37.2	1.8	13.996	X		X		X	
3	38.4	36.9	2.5	14.036	X	X	X			
4	39.2	37.1	3.1	14.011	X	X	X		X	
4	38.9	36.8	3.7	14.049	X		X	X	X	
5	39.2	36.6	5.0	14.066	X	X	X		X	X
5	39.2	36.6	5.1	14.069	X	X	X	X	X	
6	39.2	36.1	7.0	14.126	X	X	X	X	X	X

KEY:
A ='Course Attitude Positive' (academic integration factor)
B ='Course Attitude Negative' (academic integration factor)
C ='Learning' (academic integration factor)
D ='Staff/Department' (social integration factor)
E ='Friends' (social integration factor)
F ='Experience' (social integration factor)

The Best Subsets Regression above shows that with all of the factors in the model, the regression analysis explains 39.2% of the variance in Class Test score. This is a reasonable amount of variance explained by this model. We can see from the regression above that the academic integration factors explain most of the variance, 38.4% without the social integration factors, which only add 0.6% to the overall variance. Therefore it seems that the academic integration factors are the most valuable in the predictive model of exam score.



Graph 1: Plot of the Residuals

The residuals are the difference between the predicted values of the regression model and the values of the outcome observed in the sample. The straight line of this plot indicates that the residuals in this model are normally distributed, therefore this regression model is reliable.

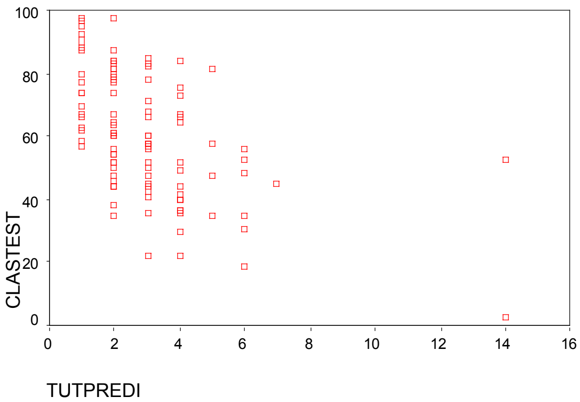
5. Analysis of Level 1 Tutor Predictions Data

The CS1P tutors were asked to predict grades for each of their students for the January CS1P exams. The purpose of this data collection was to determine whether Level 1 tutors could predict which of their students would pass or fail. Each tutor's student group is named randomly. Predicted grades were obtained from 8 of the 22 Level 1 tutors. The tutors were asked to grade their tutors from A-N in accordance with department marking schemes. These grades were then coded for analysis as follows; A = 1 B = 2 C =3 D=4 E=5 F=6 G=7 N=14 .
Note that the higher the number the lower the grade.

The Scatterplot below is the first step in the analysis of this data. It shows the general trend of the collected data and identifies outliers.

Scatterplot of classtest (%) against

tutor's predictions



This plot shows a negative correlation between class test and tutor's prediction this is what we would expect. As the class test result gets higher, the tutor's prediction gets lower (lower = higher grades given) therefore the tutor's predictions and class test score are correlated. Below is a table of the correlations between mean percent and tutor's predictions. The table gives the Pearson correlation coefficients.

Correlations

		TUTPREDI	MEAN%
TUTPREDI	Pearson Correlation	1.000	-.478**
	Sig. (2-tailed)	.	.000
	N	103	103
MEAN%	Pearson Correlation	-.478**	1.000
	Sig. (2-tailed)	.000	.
	N	103	103

**. Correlation is significant at the 0.01 level (2-tailed).

The Pearson's correlation coefficient = -.478. This indicates that there is moderately strong negative correlation between the Level 1 tutor's predictions and the students mean percentage from the January CS1P exams. The next step in this analysis is to calculate individual correlations for each tutor group to discover whether a relationship exists between tutor predictions and student exam score.

The results of the Bivariate Correlation are in the table below, The correlations were mostly of the mean percentage of lab exam and class test and tutor's predicted grade. However those tutors that predicted a grade for class test only were correlated against student's class test performance. Using departmental assessment guidelines (C = fail to progress) the tutor's predictions were coded into a pass or fail for each student, then students performance in the CS1P exams was coded into pass or fail using the same criteria. From this the percentage of correct pass or fail predictions by each individual tutor were calculated, these are displayed in the table below.

Tutor Group	Correlation coefficient- mean % vs. prediction	Correlation co-efficient- class test % vs. prediction	Sig.	% Predicted correctly (p/f)
1	-.257		.376	71%
2	-.368		.329	60%
3		-.675	.016	41%
4	-.456		.217	67%
5		-.574	.032	78%
6	.017		.960	54%
7	-.728		.002	93%
8	-.503		.040	70%

The table shows that there is a wide variance amongst the tutor predictions correlation coefficients. Four out of the eight tutor group achieved significant correlations between true exam result and predicted result. The other four groups did not have significant correlations. Tutor group 7 had the most correctly predicted pass or fail results for the students. The percentage of correct predictions in terms of the students passing or failing the January CS1P exams are quite high with most tutor groups having a percentage above 60. This would suggest that tutors could make reliable predictions about their students. The results from each of the sections above will be discussed in the section that follows.

Discussion

In this section of the paper the results will be analysed, recommendations for future research will be made, and conclusions will be drawn in regards to the project fulfilling its aims.

The aims of this project were to identify the variables that are instrumental within student failure and to build a predictive model of failure, and to follow up the findings of the study by Roddan within the same department.

Analysis of the Results

Before discussing the significant results of this study, it is important to examine the variables that were not found to be significant. Within the two questionnaires distributed such variables include the prior experience and study habit items. Roddan's results suggested that prior experience was a factor to examine in future research, having discovered that entry point score had no effect on performance. A study by Hagan and Markman (2000) found that students who have experience in at least one programming language before starting a university course perform significantly better in assessments. The results of the present study indicate that this is not the case for CS1P students. Whether or not this sample of students had; a previous computing qualification, had used a program before, had spent any hours programming, or had novice to expert computer skills, had no significant effect on class test performance. This result provides support for the view that school computing does not prepare students for university computing, despite having done computers at school it cannot be assumed that they have a substantial understanding of the subject matter. As Draper (2002) suggests the real importance of previous qualifications may be in giving the learners an accurate perception of what the subject is like, what study activities it requires and whether they would enjoy it.

Another finding from Roddan's study was that amount of revision had no clear effect on exam performance. Roddan suggested that exam performance was more likely to be due to how well they have kept up with the course and whether they understood the material. The present study therefore investigated students study habits to see whether or not these affected exam performance. None of the study habit variables had such an effect on the exam performance of this sample. This would suggest that the number of hours of independent study a CS1P student does in a typical week does not affect their exam performance. Perhaps having only collected this information on one occasion these results are not reliable. In order to discover whether students study habits effect their CS1P performance regular information on the amount of study they do would need to be collected.

Could these insignificant results be due to the instruments used? The questionnaires which measured prior experience and study habits do seem to measure what they set out to and the sample is representative of the population. Perhaps the items on the questionnaire could be improved to collect the more important aspects of studying. However the prior experience data can be trusted and as such need no further attention in future studies. The significant results will now be examined in detail.

Results from the analysis of the Understanding Question were interesting and supported Roddan's initial thought that students' self-estimates correlate with exam performance. Students' self-estimates of how well they understood the course material were highest and significant at the end of the first term in the last session. The correlations, as expected, were lowest in the first few sessions and increased quite dramatically half way through the data collection. This is explained as during the fist few weeks of the course the students are unable to assess how well they understand the material. Self-estimates correlate most highly when the students have been taught the course material and have a clearer idea of what is expected of them. This is encouraging and suggests that the students interpreted the question correctly. The results indicate that the students self-estimates begin to significantly correlate with their exam performance around week 4 of the first term. The correlations are reasonably strong ranging from 0.302 – 0.586, and they are all significant.

This would suggest that students can quite accurately judge how well they are understanding the material and that their accuracy increases with time, being highest in the last week of the first term. However we must be cautious when analyzing these results, the correlations are modest, hence, whereas some students are quite good at estimating their understanding, there are a significant number who are not. The data collection was problematic, individual student data was very intermittent and therefore analyses at this level could not be made. Despite this, the data collection was successful at identifying when students can accurately judge their own understanding.

A second significant factor is academic and social integration. It was of interest to discover whether Tinto's theory of integration (1975), applied to the CS1P students involved in this project and whether their levels of integration were predictive of exam performance. Caution must be taken when interpreting these results as the questionnaire items were completely exploratory and had not been used to test levels of academic and social integration before. Despite this, these variables provided interesting results.

Three academic and three social factors were identified through factor analysis and regression indicated that together these explained 39.2% of the variance in class test result. This is quite a low variance, however considering the novelty of the items used, this is a successful result and verifies the instrument. The regression results showed that it was the academic integration factors which explained the biggest proportion of these results (36%). Social integration did not predict a significant amount of the variance, so for CS1P students social integration is not a predictive factor, whether they integrate socially does not have a significant effect on class test result. This is very interesting and suggests that CS1P students perform differently on the class test partly due to their level of academic integration into the computing science course. More specifically the academic integration factor which accounted for the most variance was 'Learning'. This factor includes items which describe how comfortable a student is with their learning, and how well they feel they are getting on with their studies. So it seems that this is very important to CS1P students and predicts how they will do in their exams. This relates to a recent study by Rountree et al (2002), who report that students who feel confident about their learning perform best. Due to the time restrictions and the amount of variables to be analysed it was decided not to include the integration data collected from the students in the first term, however it may be interesting to look at this data and see whether levels of 'Learning' change over time and are predictive of exam result early on as with the Rountree (2002) study.

The results provide empirical support for the role of Tinto's integration model in the success or failure of the CS1P students in this course, and it can be usefully measured in practice. Research in the literature on Tinto's model is very seldom at the departmental level, this study has demonstrated that his theories are applicable at this level of investigation. Further investigation of this aspect of the CS1P student experience should be investigated, with integration information being collected more regularly since Tinto's model is designed as a longitudinal one, and those items which did not account for much variance being refined to capture more data. Furthermore, perhaps students who feel uncomfortable with their learning in CS1P should be identified and offered extra support and guidance. This could reflect on the department's expectations communicated to the students making them feel less confident about their learning.

Finally, another significant result was Level 1 tutor predictions. The majority of the tutors correctly predicted their students' pass or failure over 60% of the time. Correlations between tutor prediction and exam result were not so accurate and varied widely. However, the results indicate that the Level 1 tutor's could make reliable pass or fail judgments about their students. In practice, tutors could be asked to identify the students in their group who they perceive to be at risk of failing and offer them extra support. The sample of tutor predictions was small (8 out of 22) therefore it is not clear whether the success of these predictions is representative of all of the CS1P tutors.

Despite this, this aspect of research was worth pursuing and has highlighted the potential for tutors to be involved in identifying students at risk of failing the CS1P course, which would be simple enough for the department to put into practice. I would suggest that this aspect be measured again with a larger, more representative sample of tutors, and on more than one occasion, it maybe interesting to discover when the tutors can make accurate predictions. This variable, could be used in conjunction with other measures to help to identify those students who are at risk, and to intervene.

Possible Confounds to this Study

This academic year the tariff to get into the University of Glasgow computing science course was increased, this may have affected the results. It will be interesting to discover what affects this has on the success rate at the end of the university year. Also during this study, other interventions were involving students in an attempt to increase student success, one of these was APASS a resource whereby first year students meet with students from higher years to ask questions about aspects of the course they don't understand. This is unlikely to have affected my results directly, but it will be interesting to see whether it affects the June exams at the end of the year. It would also be interesting to discover whether the students attending APASS have higher levels of academic integration and specifically 'Learning' as measured by this study which had the highest predictive power of class test.

Significance of the Study

The significance of this study has been in developing and implementing exploratory instruments to measure potential influences on student attrition that had not previously been measured in Computing Science. The Tinto integration questionnaires were implemented with relative success and encourages more investigation. Also the 'Understanding Question' was an exploratory investigation following up the lead from Roddan's study. This project shows that this variable was worth pursuing and is worth pursuing in the future, developing a more successful way to gather regular data for each student , perhaps gathering data in tutorials, which students attend more regularly. The merit of this project also has been in identifying variables which need no further investigation, such as prior experience, these variables were not significant and are unlikely to have a consequence on students' performance.

Furthermore, this study has highlighted that general theories do apply to this Computing Science course, it may not be necessary to design experiments which look for factors relevant to programming, it may be more likely that the same factors determining drop-out at the university level effect drop-out in Computing Science as well, but perhaps at different times, to different extents and different combinations. This is a valuable point and one which other research into computing science should keep in mind, not simply to focus on factors of programming, as with such literature as Gilmore (2002). Importantly, this study obtained a significant amount of data from the CS1P students, data which could be analysed in terms of the June exams, and analysed in more depth with additional information, such as APASS, to provide more clues of the important influences on CS1P student progression into second year.

Conclusion

It is impossible to discover exactly what factors equal failure for a student, and a great deal of the variability in achievement in Computing Science courses remains unaccounted for. This project has been beneficial in identifying aspects for further examination which will lead on from where this project has ended.

The Experience of Being a Computing Science Student

Please enter your matriculation number here:							

1. Age: _______________
2. What is your intended degree? ______________________________________

3. Regarding **Computing Science** please indicate in the table below your most advanced qualifications (if any):

Subject: (e.g. computer studies/ information systems/ IT.)	**Type of Qualification:** (e.g. A-level, SYS, Higher, Standard Grades, HNC.)	**Grade Obtained:**

5. If you have done programming, *other than as part of the courses above*, then how much total time do you think you have spent computer programming?
   - ❒ 0 hours
   - ❒ 10 hours
   - ❒ 100 hours (e.g. 10 hours per week for 10 weeks)
   - ❒ 1000 hours (e.g. 40 hours per week for 25 weeks)
   - ❒ 10,000 + hours

6. If you have any experience programming, please indicate the programming language(s) you have used (e.g. JAVA).

____

7. Do you have a computer at home? (please tick a box) ❒ Yes ❒ No

8. Tick the box if you know how to do the following on either Windows or Macintosh computers.

- ☐ Open and close folders
- ☐ Move a document from one folder to another
- ☐ Copy documents to and from floppy disks
- ☐ Start up an application
- ☐ Shut down the computer
- ☐ Open and Close Microsoft Word documents
- ☐ Save Microsoft Word documents
- ☐ Print Microsoft Word documents
- ☐ Use Cut, Copy and Paste in Microsoft Word
- ☐ Format paragraphs in Microsoft Word
- ☐ Use page headers and footers in Microsoft Word
- ☐ Number pages in Microsoft Word
- ☐ Use Styles in Microsoft Word
- ☐ Use Tables in Microsoft Word
- ☐ Save Microsoft Word documents in various formats

9. On the scale provided below, please rate your general computer skills and expertise by circling the appropriate number.

Novice						Expert
1	2	3	4	5	6	7

10. The following questions should be answered quickly, without careful thought, to give an impression of your feelings at the moment.

Please circle the number closest to your opinion. If the question is Not Applicable, tick "N/A".

	Strongly Disagree						Strongly Agree	N/A
Studying Computing Science is useful	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
This course demands things of me that I don't like (e.g. labs, maths)	-3	-2	-1	0	+1	+2	+3	❐
University is not providing me with some of the kinds of conversations I want	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
Getting a good grade in this course is important to me	-3	-2	-1	0	+1	+2	+3	❐
This course prevents me from engaging in learning activities that I like (e.g. essay writing, group work)	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I like the learning activities that I am asked to complete in this course (eg. tutorials, programming exercises)	-3	-2	-1	0	+1	+2	+3	❐
I would like to get to know staff at the University	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
This course blocks me from doing things which are important to my learning (e.g. have questions answered, have time to think before the next thing is presented)	-3	-2	-1	0	+1	+2	+3	❐
I find Computing Science interesting	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
So far I am understanding the material as well as I want to	-3	-2	-1	0	+1	+2	+3	❐
I have the skills to take effective notes in lectures	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I feel comfortable being a student at this University	-3	-2	-1	0	+1	+2	+3	❐
Apart from getting a qualification there is no value in coming to this University	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
The other students at the University are not worth getting to know	-3	-2	-1	0	+1	+2	+3	❐
My preferred kinds of socialising (e.g. clubbing, hill walking, dinner parties) do not fit student life	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
Getting to know students and staff is useful to me	-3	-2	-1	0	+1	+2	+3	❐
I feel I know how to make friends at University	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I like the conversations I find myself having with students at University	-3	-2	-1	0	+1	+2	+3	❐
I like the conversations I find myself having with staff at University	-3	-2	-1	0	+1	+2	+3	❐
I enjoy the social activities that other students suggest	-3	-2	-1	0	+1	+2	+3	❐
I feel I know how to talk to other students	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I feel comfortable around campus, the departments, in lectures etc.	-3	-2	-1	0	+1	+2	+3	❐
Studying Computing will lead to the job or career that I want	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I feel comfortable approaching staff whenever I feel I need to	-3	-2	-1	0	+1	+2	+3	❐
I fit in with other students in the Computing Science class	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
When I'm with other people outside the University I feel embarrassed that I am a student	-3	-2	-1	0	+1	+2	+3	❐
Going to this University fits with the kind of person I want to be	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I feel comfortable being a student in the UK today	-3	-2	-1	0	+1	+2	+3	❐

I want to get on with people outside the University	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
People outside the University tend to accept my going there as worthwhile	-3	-2	-1	0	+1	+2	+3	❐
Being at the University of Glasgow is impressive to others	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I fit in with other students in the University	-3	-2	-1	0	+1	+2	+3	❐
I feel better about myself being a student than if I were doing something else	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐

	Strongly Disagree						Strongly Agree	N/A
Going to University and being a student helps me get on with other people such as family, friends and employers	-3	-2	-1	0	+1	+2	+3	❐
Going to University makes me fit in better in life outside the University	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	❐
I feel comfortable telling others I go to the University to Glasgow	-3	-2	-1	0	+1	+2	+3	❐

11. How many staff have you had a personal interaction with (however brief)? _________ (please give a number)

12. How many students at Glasgow University do you know? _________ (please give a number)

13. Please indicate the degree to which the following statements are true of you right now. Please circle your response and refer only to your Computing Science course.

	Not at all true of me						Very true of me
I often think to myself "What if I do badly in Computing Science?"	**1**	**2**	**3**	**4**	**5**	**6**	**7**
I want to learn as much as possible from this course	1	2	3	4	5	6	7
It is important to me to do better than the other students in this course	**1**	**2**	**3**	**4**	**5**	**6**	**7**
My goal in Computing Science is to get a better grade than most of the students	1	2	3	4	5	6	7
I want to do as little as possible in this course	**1**	**2**	**3**	**4**	**5**	**6**	**7**
I worry about the possibility of getting a bad grade in Computing Science	1	2	3	4	5	6	7
It is important for me to understand the content of this course as thoroughly as possible	**1**	**2**	**3**	**4**	**5**	**6**	**7**
I just want to do what I am supposed to do in this course and get it done	1	2	3	4	5	6	7
I am striving to demonstrate my ability relative to others in this course	**1**	**2**	**3**	**4**	**5**	**6**	**7**
My fear of performing poorly in this course is often what motivates me	1	2	3	4	5	6	7
I hope to have gained a broader and deeper knowledge of Computing Science when I am done with this course	**1**	**2**	**3**	**4**	**5**	**6**	**7**
I am motivated by the thought of out performing my peers in this course	1	2	3	4	5	6	7
I want to do things as easily as possible in this class so that I wont have to work too hard	**1**	**2**	**3**	**4**	**5**	**6**	**7**
I just want to avoid doing poorly in Computing Science	1	2	3	4	5	6	7
I desire to completely master the materials presented in this course	**1**	**2**	**3**	**4**	**5**	**6**	**7**
It is important to me to do well compared to others in Computing Science	1	2	3	4	5	6	7
I'm afraid that if I ask the teaching staff or my instructors dumb questions, they might not think I am very smart	**1**	**2**	**3**	**4**	**5**	**6**	**7**
In a course like this, I prefer material that arouses my curiosity, even if it is difficult to learn	1	2	3	4	5	6	7

I want to do well in Computing Science to show my ability to my family friends, advisers or others	**1**	**2**	**3**	**4**	**5**	**6**	**7**
My goal for this course is to avoid performing poorly	1	2	3	4	5	6	7
In Computing Science I prefer course material that really challenges me so I can learn new things.	**1**	**2**	**3**	**4**	**5**	**6**	**7**

14. If you have time, please answer these questions as well:

(i) What are your reasons for studying computing science?

(ii) What are the best and worst aspects of getting to know staff and students here?

(iii) What is good and bad about presenting yourself to outsiders as someone from the University of Glasgow?

Thank you for taking the time to complete this questionnaire

References

1) Elkins, S.A. , Braxton J.M., and James G.W. (2000). *Tinto's Separation Stage and its Influence on First-Semester College Student Experience*. Research in Higher Education, Vol.41, No. 2.

2) Guarino, A., Michael W.B., Hocevar D. (1998). *Self-Monitoring and Student Integration of Community College Students*. The Journal of Social Psychology, 138, (6), 754-757.

3) Bray, N.J., Braxton, J.M., Sullivan, Anna Shaw. (1999). *The Influence of Stress-Related Coping Strategies on College Student Departure Decisions*. Journal of College Student Development, November/December Vol. 40, No.6.

4) Gracia, L., and Jenkins, E. (2002) *An Exploration of Student Failure on an Undergraduate Accounting Programme of Study*. Accounting Education 11, (1), 93-107.

5) Jenkins, T. (2002). *On the Difficulty of Learning to Program*.
http://www.psy.gla.ac.uk/~steve/localed/jenkins.html

6) Stephens, D. and Creaser, C. (2002). *Information Science student IT experience and attitude toward computers: results of a five-year longitudinal study*. Innovations in Teaching And Learning in Information and Computer Sciences.
http://www.ics.Itsn.ac.uk/pub/italics/issue2/dstephens_b/0005.html

7) Thomas, S.L., (2000) *Ties That Bind: A Social Network Approach to Understanding Student Integration and Persistence*. The Journal of Higher Education, Vol. 71, No.5 (September/October).

8) Hinett, K. (1998). *The Role of Dialogue and Self Assessment in Improving Student Learning*. BERA Annual Conference, The Queen's University of Belfast, August 27-30, 1998.

9) Jenkins, T., Davy, J. (2001). *Diversity and Motivation in Introductory Programming*. Proceedings of ITiCSE 2001, pp. 53-56.

10) Mooney, T. (2002). *The Feet That Vote*. http://www.EducationGuardian.co.uk

11) Braxton, J.M., Milem, J.F., Sullivan, Anna Shaw. *The Influence of Active Learning on the College Student Departure Process*. The Journal of Higher Education, Vol. 71, No.5, (September/October).

12) Palmer, J. (2001). *Student Drop-out: A case study in new managerialist policy*. Journal of Further and Higher Education, Vol.25, No.3.

13) Torres, J.B., Solberg, V.S. (2001). *Role of Self-Efficacy, Stress, Social Integration, and Family Support in Latino College Student Persistence and Health*. Journal of Vocational Behaviour 59, 53-63.

14) Montmarquette, C., Mahseredjian, S., Houle R. (2001). *The Determinants of university dropouts: a bivariate probability model with sample selection*. Economics of Education Review, 20, 475-484.

15) Borglum, K., Kubala, T. (2000). *Academic and Social Integration of Community College Students: A Case Study*. Community College Journal of Research and Practice, 24: 577-586.

16) Glossop, C. (2002). *Student nurse attrition: use of an exit-interview procedure to determine students' leaving reasons*. Nurse Education Today, 22, 375-386.

17) Mannan, Muhammad, Abdul. (2001). *An assessment of the academic and social integration as percieved by the students in the University of Papua New Guinea*. Higher Education, 41, 283-298.

18) Norusis, M.J. (2000). *Guide to Data Analysis*. Prentice Hall.

19) Field, A. (2002). *Discovering Statistics using SPSS for Windows*. Sage publications.

20) Gilmore, K. (2002). *Predicting Level One Programming Students at Risk*. Final Year Thesis, Department of Computing Science, University of Glasgow.

21) Roddan, M. (2002). *The Determinants of Student Failure and Attrition in First Year Computing Science*. Final Year Thesis, Department of Computing Science, University of Glasgow.

22) Rountree, N., Rountree, J., and Robins, A. (2002). *Predictors of Success and Failure in a CS1 Course*. SIGCSE Bulletin, Vol. 34, No.4.

23) Tinto, V. (1982). *Limits of Theory and Practice in Student Attrition.* Journal of Higher Education, Vol. 53, Issue 6, 687-700.

24) Tinto, V. (1988). *Stages on Student Departure: Reflections on the Longitudinal Character of Student Leaving*. Journal of Higher Education, Vol. 59, Issue 4, 438-455.

25) Tinto, V. (1975). *Dropout from Higher Education: A Theoretical Synthesis of Recent Research.* Review of Educational Research, Vol. 45, pp. 89-125.

26) Patrick, B. (2001). Students Matter: Student Retention: who stays and who leaves? The University Newsletter, http://www.gla.ac.uk/newsletter/226/html/news29.html.