Computer Science Education
Vol. 00, No. 0, Month 2007, pp. 1–19

Optimists Have More Fun, But Do They Learn Better? On the influence of emotional and social factors on learning introductory computer science

Jens Bennedsen[a]* and Michael E. Caspersen[b]
[a]*IT University West, Denmark;* [b]*University of Aarhus, Denmark*

In order to better understand predictors of success and, when possible, improve the design of the first year computer science courses at university to increase the likelihood of success, we study a number of factors that may potentially indicate students' computer science aptitude. Based on findings in general education, we have studied the influence of emotional and social factors on students' learning outcomes in introductory computer science courses. Emotional health and social well-being have been measured in terms of five variables: perfectionism, self-esteem, coping tactics, affective states and optimism. Surprisingly, we found no correlation between emotional health and social well-being on the one hand and success in computer science as indicated by course grades on the other. However, in most of the courses, the students who pass have a statistically significant higher self-esteem than those who do not. Our hypothesis was that there would be a positive correlation between emotional and social factors and success in computer science as indicated by the course grade, since others have found the variables perfectionism, self-esteem and affective states to be predictors of success. We identify potential explanations for this seeming contradiction.

Introduction

Currently, there is a strong focus on the declining number of students graduating in computer science (Denning, 2004; Madsen, 2006). An indication of this strong focus is that ACM's Education Board recently decided to establish a task force with the focus of increasing enrolment in computer science. Changing the state of affairs is indeed a challenge, but the challenge has two facets: improving recruitment and increasing the number of students that succeed in computer science. The focus of this study is on factors that may influence the students' success rate. In this context we are

*Corresponding author. IT University West, Fuglesangs Allé 20, DK-8210 Aarhus V, Denmark. E-mail: jbb@it-vest.dk

ISSN 0899-3408 (print)/ISSN 1744-5175 (online)/07/000001-19

DOI: 10.1080/08993400701791133

specifically interested in the influence of emotional and social factors on learning introductory computer science. If we can find factors that are indicators of success, we might have a chance of improving or acting upon these factors and hence increase the number of students that succeed in computer science.

Introductory computer science is tightly coupled with programming which is notoriously considered to be difficult. For almost 40 years, teaching programming to novices has been considered a big challenge, and it still is (Dijkstra, 1969; Gries, 1974; Soloway & Spohrer, 1989; Tucker, 1996; McCracken et al., 2001; Robins et al., 2003); in fact it is considered one of seven grand challenges in computing education (McGettrick et al., 2005).

The question we pose is: What can be done to improve the freshman year such that a greater number of students learn more and become better equipped to cope with the following years of education?

In general, we can do several things to (try to) improve the state of affairs. We can improve *what* we teach (content), *how* we teach (pedagogical design), the students' *learning conditions* or the *teaching milieu*. Hiim and Hippe (2006) have developed a didactical model in which they identify six aspects that influence students' learning outcomes:

1. *Student learning premises:* knowledge, experiences, attitudes, and skills that the students already possess when they come for the first lesson of the course.
2. *External conditions:* conditions that limit or make learning possible, such as equipment, artefacts, time, place, classroom settings, teacher's resources, learning resources, etc.
3. *Objectives for the learning activity:* what the students should learn from the course/activity in terms of knowledge, skills, attitudes and competencies.
4. *Contents:* what the course is about, how content is selected, adjusted and presented.
5. *Learning process:* the process of change within the learning subjects (the students), and reflections on how the intended changes are facilitated.
6. *Evaluation:* assessment or evaluation in relation to the teaching process, in relation to the objectives for the course, and in relation to the students' learning (Hiim & Hippe, 2006, pp. 28–30).

Three of these aspects are concerned with what we teach (3, 4 and 6), one is concerned with how we teach (5), one is concerned with the students' own learning conditions (1), and one is concerned with the teaching milieu (2). This study focuses on students' personal learning premises. To help focus our resources, we look for factors that are predictors of success. In this research, we study the influence of emotional and social factors on students' learning outcomes. If these factors turn out to be predictors of success, we should use this knowledge either directly or indirectly depending on how stable the personality traits are, and focus on improving the pedagogical design and general study environment to help these students. One such initiative could be to change the teaching format from large group lecturing to class-based teaching for smaller groups in order to foster a closer and more caring relationship between the students and the teacher; Argyle (1994) showed that this enhances students' self-esteem. Another initiative could be to improve feedback and make exemplary solutions to assignments available such that students with a high demand for perfection know in more detail what is required for a top grade. Other ideas could be adopted from Lawrence's book (1996) on enhancing the self-esteem in the classroom (even though it is aimed for primary school).

Related work on predictors of success

Evans and Simkin (1989) sum up the arguments given in many studies for performing this kind of study (p. 1322):

1. Discriminating among enrolment applicants.
2. Advising students on their majors.
3. Identifying productive programmers.
4. Identifying employees who might best profit from additional training.
5. Improving computer classes for non-CIS majors.
6. Determining the importance of oft-cited predictors of computer competency, such as gender or math ability.
7. Exploring the relationship between programming abilities and other cognitive reasoning processes.

Evans and Simkin's study was carried out during a time when many students enrolled in computer science classes. This is not the case today, and so the first argument is not currently relevant. Evans and Simkin mention 'improving computer classes for non-CIS majors'. Our intention is to find factors that may improve computer classes for CS majors in particular.

A substantial amount of research has been conducted in order to identify variables that are predictors of success for students aiming for a university degree—variables that account for parts of the students' learning premises. The investigated variables encompass among other things gender (Rountree et al., 2004; Bennedsen & Caspersen, 2005; Ventura, 2005), the educational level of parents (Ting & Robinson, 1998), ACT/SAT scores[1] (Butcher & Muth, 1985; Sanders, 1998; Bennedsen & Caspersen, 2005; Bennedsen & Caspersen, 2006), performance in prior courses (Chamillard, 2006), and abstraction ability (Sprague & Schahczenski, 2002; Bennedsen & Caspersen, 2006). All of the above studies have focused on (academic) knowledge, but the results are not conclusive and point in different directions—e.g., Ventura (2005) concludes that math is not a predictor of success for an introductory object-oriented programming course while Bennedsen and Caspersen (2005) concludes the opposite. The general pattern is a rather weak (if any) impact of the analysed variables. In general, the studies show that gender and abstraction ability have no influence, previous math scores have an impact in some cases but not in others, and scores in previous courses can be used to predict students' outcome in following courses.

Mouw and Khanna (1993) conclude from a review of 39 studies that only 5% to 25% of the variance in college performance can be accounted for by aptitude scores such as the ACT and the SAT. The lack of clear and definite conclusions has caused researchers to study the impact of other types of factors. Szulecka et al. (1987) have suggested that the major causes of attrition among first year college students are due to emotional rather than academic factors. This is consistent with the conclusions of Pascarella and Terenzini (1991) from a comprehensive literature review: 'Thus, as a measure of successful adaptation to an academic environment, grades tend to reflect not only requisite intellectual skills but also desirable personal work habits and attitudes' (p. 388).

Many psychological variables affect the success and retention of students in an educational setting. Brooks and DuBois (1995) found that emotional variables have a strong influence on how well students adjusted to their first year at college. This is a strong predictor of academic success (Van Heyningen, 1997). Furthermore, Leafgran (1989) claims that emotional health has a positive influence on students' success in college. Pritchard and Wilson (2003) conclude that social factors have a positive influence on student performance. High self-confidence, self-control and a conscientious personality are associated with higher academic performance (Wiggins & Blackburn, 1969; Wolfe & Johnson, 1995). In addition, students who are perfectionists tend to adjust better to college and, as a result, show higher rates of retention (Rice & Mirzadeh, 2000).

Students and study programs

The students in this study were freshmen in computer science at the University of Aarhus during the academic year 2005–2006. We investigated the correlation between emotional health and social well-being on the one hand and success in selected first-year courses in computer science on the other.

The general structure of an academic year at the Faculty of Science at the University of Aarhus is four quarters (each of seven weeks), each followed by a two- to four-week examination period. Students take three courses (each of five ECTS[2]) in each quarter. The students need to pass individual exams of all courses in order to pass the first year program in computer science. Some of the courses require that the students have passed an exam in a previous course (e.g., 'Algorithms & data structures 2' requires that the student have passed 'Algorithms & data structures 1' which in turn requires that the student have passed 'Introduction to programming'). Therefore, there will typically be fewer students participating in the later courses than in the first courses. The drop out rate in the first year is approximately 20%.

Table 1 presents the first year program in computer science.

The courses in italics are those with a grading scheme specific enough to be analysed. 'Calculus 2' is a math course that is required for the computer science majors to take. Since the study focuses solely on computer science, 'Calculus 2' is excluded from the analysis.

Table 2 gives a short description of the content of each course included in this study. Course descriptions can be found at www.daimi.au.dk/courses/descriptions/introductory-courses/. The code in parentheses after the course name indicates the corresponding course in CC2001 (Engel & Roberts, 2001).

Research method

In this section, we discuss the methodology used in identifying predictors of success for the courses described in the previous section. We then outline the research hypothesis, describe the data and how they were provided, and provide details on the subjects involved in the study. After presenting and discussing the operationalization

Table 1. First-year program in computer science

1.	Introduction to programming	Perspectives on CS	Calculus 1
2.	*Programming 2*	*Usability*	*Calculus 2*
3.	Algorithms & data structures 1	Web technology	*Computer architecture*
4.	*Algorithms & data structures 2*	*Programming languages*	*Regularity & automata*

Table 2. Description of the courses

Course	Content
Programming 2 (CS111$_{O}$)	Language concepts (polymorphism, events, exceptions, streams, and threads); program design; (recursive) data structures; class hierarchies; frameworks
Usability (CS350)	Human machine interaction; UI components; interaction; UI tools; usability methods
Computer architecture (CS220)	Hierarchical computer architecture (digital level, micro architecture level, conventional level, and OS level); assembly language; hardware architecture; external devices
Algorithms and data structures 2 (CS210$_{T}$)	Algorithmic paradigms (divide and conquer, dynamic programming, greedy algorithms); graph algorithms (graph traversal, topological sorting, spanning trees, shortest paths, transitive closure); text processing (pattern matching, tries, text compression, text similarity)
Programming languages (CS344 and CS345)	Functional programming (higher order functions, lazy evaluation, polymorphism, modules); logic programming (unification, back-tracking, knowledge representation, logic grammars)
Regularity & automata (–)	Formal models of regularity (finite automata, regular expressions, regular grammars); proof techniques (invariance, structural induction); applications in CS

of the hypothesis and the test instrument we have applied, we present the statistical analysis.

Hypothesis

We will study the influence of emotional and social factors on a student's learning outcome; our hypothesis is:

- There is a positive correlation between a student's well-being and emotional health and his or her success in introductory computer science courses as defined by the course grades.

Data

Two data sources will be used in this study. Information comes from the administrative system at the university (exam scores) and a questionnaire with questions about social well-being, emotional health, motivation, etc. The questionnaire was answered by the students during week six and seven of their introductory year (i.e., at the end of the first quarter).

The result of the final exam is used as an indicator for success—higher grade, more success. Two scales of marks are used: pass/fail and a ten-ary scale (Exam-scale, n.d.). The grades of the Danish ten-ary scale are: 00, 03, 5, 6, 7, 8, 9, 10, 11 and13. A student needs at least a '6' to pass an exam. Table 3 shows the official conversion from the Danish scale of marks to the ECTS scale of marks (European Union, 2006). Unfortunately, an official conversion to the North American scale does not exist; we have included our conversion.

Courses with binary grading (pass/fail) do not provide a normal distribution of grades and hence do not allow for the kind of statistical analysis we are aiming for; we therefore only consider courses using the ten-ary scale of marks.

In general, the result of the grade that a student gets is solely determined by the final exam. In several of the courses, the student need to pass mandatory assignments as a prerequisite for the final exam; the grades of these mandatory assignments are not included in the final exam score.

Table 3. Conversion from Danish scale of marks

Danish scale of marks	ECTS scale of marks	North American scale of marks
11 and 13	A	A
10	B	B
8 and 9	C	C
7	D	C
6	E	D
03 and 5	Fx	F
00	F	F

Participants

One hundred thirty-four computer science students participated in one or more of the exams of the courses in the first year program in computer science (Table 1). However, some of the students participated only in the courses in the first quarter, so the total population in the study is 100.

Two hundred and six students voluntarily answered the questionnaire (some from study programs other than computer science); 77 have participated in one or more of the computer science exams. Table 4 summarizes the number of students within each course who participated in the exam, the number of students who participated in the exams who also answered the questionnaire, and the response rate. As can be seen from the table, the average response rate is rather high except for the course 'Programming languages' which therefore is excluded from our analysis.

To check for representativity of the students who have answered the questionnaire, we have performed Mann-Whitney-U tests. We compared the exam scores of the students who answered the questionnaire with the exam results of the students who did not answer the questionnaire. Since it is not the same group of students who have participated in all the exams, we have performed a test for each of the courses. The students who have answered the questionnaire are representative in all the courses but 'Algorithms & data structures 2'; in this course, the students who answered the questionnaire had a significantly higher exam score.

The test instrument

As our hypothesis reflects, we focus our attention on two non-academic factors: emotional health and social well-being. We use five variables to measure students' emotional health and social well-being: perfectionism (PERFECT), self-esteem (SELFEST), coping tactics (COPE), affective states (POM) and optimism (OPTIMISM). The choice of factors is motivated by many studies in general educational research (Raad & Schouwenburg, 1996; Pritchard & Wilson, 2003).

Several tests have been produced and validated in the field of (educational) psychology. In our study, we will use a battery of such tests and evaluate the

Table 4. Response rate courses

Course	Number of students who participated in exam	Number of students who answered questionnaire	Response rate
Programming 2	76	61	80%
Usability	100	70	70%
Computer architecture	87	56	64%
Algorithms and data structures 2	83	49	59%
Programming languages	40	9	23%
Regularity & automata	70	46	66%

correlation with the exam results. Perfectionism is assessed using a subscale of the Eating Disorders Inventory (EDI) (Garner et al., 1983). One of the hypotheses regarding eating disorders is that there is a correlation between an eating disorder and perfectionism. We use the part of the scale that addresses perfectionism. Students will respond to statements about their performance levels in activities and the influence of the expectations of others (e.g., family, teachers, parents), such as, 'Only outstanding performance is good enough in my family'. Responses indicate the participant's agreement based on a 6-point scale ranging from '1' (never) to '6' (always). The EDI has been shown to be valid and reliable (Garner et al., 1983). Lee et al. (1998) evaluated the cross-cultural validity of the Chinese version of the EDI and found that the profiles of Chinese and Canadian patients were similar to each other.

Self-esteem is measured using the Rosenberg Self-Esteem Scale (Rosenberg, 1965). This scale is probably the most widely used to measure self-esteem. The scale has 10 questions that address personal feelings plus positive and negative emotions (e.g., 'I feel I have a number of good qualities'). Students responded on a 4-point scale ranging from 'strongly agree' to 'strongly disagree'. This scale has been shown to be reliable and valid in many populations (Bosson et al., 2000) and has recently been used among 53 nations from different cultures (Schmitt & Allik, 2005). Schmitt and Allik (2005) found 'the RSES [Rosenberg Self-Esteem Scale] factor structure was largely invariant across nations' (p. 623); their finding supports cross-cultural equivalence of the scale.

Coping tactics are measured via the brief COPE (Carver, 1997). This 28-item questionnaire contains 14 tactics (e.g., seeking emotional support, giving up, etc). Students responded to how they would deal with a stressful event on a 4-point Likert scale (Corbetta, 2003 p. 170ff.) ranging from 'I wouldn't do this at all' to 'I would do this a lot'. This measure has been tested in a variety of populations, and the measure has been validated and shown to be reliable (Carver, 1997). Affective states are measured using the 30-item version of the Profile of Mood States (McNair et al., 1981). This Likert-type questionnaire assesses the mood states of tension, depression, anger, vigour, fatigue and confusion. This measure has been tested on several populations and has been shown to be reliable and valid (Shin & Colling, 2000). In particular, Yeun and Shin-Park (2006) have found that the instrument can be used both in the US and Korea.

Optimism is assessed via the Defensive Pessimism Scale (Norem & Cantor, 1986). The students will indicate the degree to which each of 11 statements describing characteristics of either optimism or pessimism is representative of their thoughts and behaviour in academic situations. Previous studies that used this questionnaire have found this scale to be reliable and valid (Sanna, 1998).

All of the above instruments are originally in English and intended to be used in the US. Van de Vijver and Hambleton (1996) discuss three types of bias in test translation and application in a different culture:

- *Construct bias*. The construct shows non-negligible differences across cultures.
- *Method bias*. Validity-threatening factors related to instrument administration.

- *Item bias*. Problems at the item level such as poor wording or inappropriateness in a given culture.

We expect the Danish and American cultures to be largely the same for the five concepts and therefore ignore the possibility of construct bias. Furthermore, we administer the test in a similar way as, e.g., Pritchard and Wilson (2003). Consequently, we disregard the possibility of method bias. In our study, we addressed Danish students and in order to ensure that differences in the participants' English competences did not influence the result, we translated the instruments from English to Danish. We therefore need to ensure that the Danish instrument measures the same as the original English one. Weeks et al. (2007) identified five test translation methods: one-way translations, back translations, bilingual techniques (a bilingual person takes the test in both known languages), committee approach and pre-test procedures. We used a one-way translation procedure (one of the authors translated the instruments; when in doubt, he discussed the translation with fellow researchers). After the translation, both the Danish and English test was taken by the same person and the Danish version adjusted according to the feedback that was provided (bilingual technique). English is a more verbose language than Danish. Consequently, some of the words in the Profile of Moods Scale translated to the same Danish word; therefore the Danish translation of this scale consisted of 26 items as opposed to 30 items in the English version.

Statistical analysis

The goal is to find how much impact (if any) the five variables have on a student's success (i.e., the result of the examination). One way to obtain this answer is to use a multiple regression analysis based on a model that is as simple as possible using the variables in question and the relevant interaction variables (i.e., combination of the variables).

We end up with a linear regression model describing the functional relationship between a predicted variable (exam score) and a set of predictor variables. The linear regression model has the form: $\lambda = \alpha + \beta_1\chi_1 + \ldots + \beta_n\chi_n$, where λ is the predicted variable, α is a constant displacement, and each χ_i is a predictor variable with corresponding coefficient β_i.

In order to use the multiple regression model, five prerequisites need to be fulfilled: linearity; normal distribution; homoscedasticity (the conditional distribution has constant standard deviation throughout the range of values of the explanatory variables); no collinearity (two or more variables have a strong linear relationship, i.e., explains the same); and no problematic outliers (an observation falls far from the rest of the data and the mean is highly influenced). Scatter-plot of the data indicates that the requirements are met. Shapiro-Wilks W tests for the variables verify that their distribution can be described as normal (they are all very close to '1'). Levene's tests was performed for the variables and showed homogeneity of variances. Pearson correlation tests show that none of the variables have strong linear relationships.

The linear regression analysis is performed by starting with a model consisting of all five predictor variables and then iteratively removing a predictor variable that is not statistically significant at the 5% significant level (in the case of more than one variable being not statistically significant, we remove the one that is least significant).

Before performing this type of empirical quantitative research, it is important to consider the required sample size so that the risk of rejecting a hypothesis when it actually is true (called a type II error, denoted ß), or the other way around (called a type I error, denoted α) is acceptable. In order to estimate the sample size, we consider three concepts (Cohen, 2003):

- *Effect size* (f^2), or the salience of the treatment relative to the noise in measurement.
- *Alpha level* (α, or confidence level), or the odds that the observed result is due to chance.
- *Power* (1-ß), or the odds that you will observe a treatment effect when it occurs.

Having $\alpha = 0.05$, 1-ß = 0.8 and $f^2 = 0.15$ and five predictor variables, we can calculate the required sample size to be 91 (Soper, 2007). Given that there are over 100 computer science students, it seems reasonable to perform the statistical analysis.

The analysis of the data was performed in Microsoft Excel using the add-in for 'data analysis' and in SPSS version 13.

Results

This section describes the results of the statistical analysis. In general, the correlation between the five predictor variables and the outcome of the exams depends at most on one variable using a 5% confidence interval.

For the courses 'Programming 2' and 'Algorithm & data structures 2', no variables correlate with the students' exam score; the courses are therefore excluded from Table 5.

The exclusion of PERFECT, SELFEST and POM from Table 5 indicates that these variables do not predict success in any of the courses. COPE predicts only one course, and that is a very weak prediction ($r = 0.233$). Furthermore, the *p*-value is just above 5%.

The variable that predicts the best is OPTIMISM; but even the strongest prediction is weak: the correlation (Pearson correlation coefficient, r) between 'Regularity & automata' and OPTIMISM is only 0.342 (0.5 is traditionally considered to be the threshold value for correlation when considering only one variable. However, as Cohen (1988) notes, such a threshold is more or less arbitrary. In general psychological research he suggests that an r-value in the interval 0.3 – 0.5 is a medium correlation and above 0.5 is a strong correlation). The correlation between OPTIMISM and the exam score in 'Regularity & automata' is illustrated in Figure 1.

Table 5. Predictor variables for courses

Course	Correlated variable	*p*-value	*N*	Correlation (r)
Usability	COPE	5.2%	71	0.233
Computer architecture	OPTIMISM	3.5%	57	0.281
Regularity & automata	OPTIMISM	1.9%	47	0.342

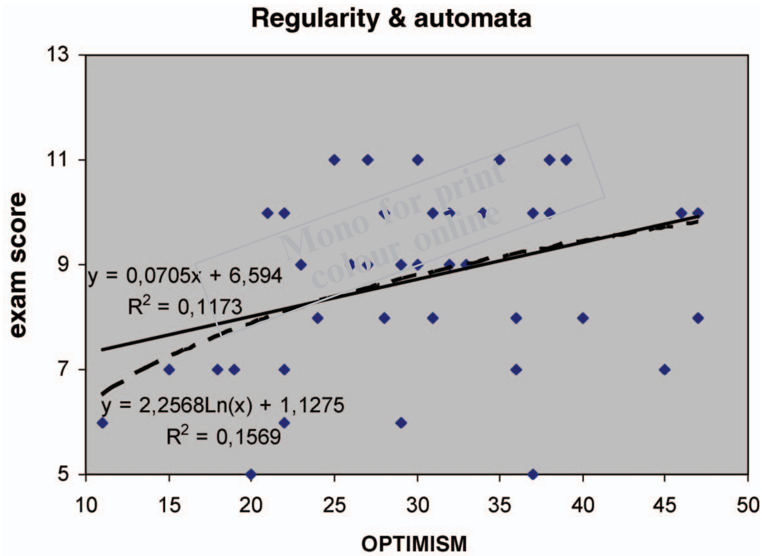

Figure 1. Scatter plot of 'Regularity & automata' and OPTIMISM (dashed line = logarithmic trendline)

The correlation between OPTIMISM and 'Regularity & automata' can be raised using a logarithmic regression; the correlation coefficient is raised to 0.396 (Figure 1). The other courses do not show the same logarithmic correlation.

Having lower sample size than initially expected raises the question of the statistical power of the results. In other words, given the results, what are the odds of confirming our hypothesis correctly? We have calculated the power given the observed correlation, number of students and significance. The results are presented in Table 6. In all of the four analyses, the chance of confirming the correlation between the variables is low.

Previously we have focused on students completing the courses. Another very relevant group is 'non-completers'—do they have different well-being and emotional health than completers? Ninety-two computer science majors enrolled in 2005. 71 of the 92 enrolled students answered our questionnaire. Of the enrolled students, only six passed all twelve courses; it is very common for the students not to pass all exams in their first year. All of the 92 students enrolled in 2005 participated in one or more of the 12 exams. 81 of the 92 enrolled students have passed one or more of the exams involved in this study and 87 students have participated in one or more of the exams involved in this study. The term 'non-completers' is therefore not precise. Consequently, we have looked for differences between the students who have participated in the exam for the 10 computer science courses and those who have not. We have used a Mann-Whitney-U test with a significance level of 5%. The result is presented in Table 7.

Table 6. The statistical power of the correlations

Course	Correlated variable	Power (1-ß)
Usability	COPE	53%
Computer architecture	OPTIMISM	53%
Regularity & automata	OPTIMISM	54%
Regularity & automata	OPTIMISM (logarithmic correlation)	71%

Table 7. Factors that differ between passing and not passing students

Course	Number of participants who have answered the questionnaire and **passed** the exam	Number of participants who have answered the questionnaire and **not passed** the exam	Factors that differ
Introduction to programming	65	6	Too few non-passers to run test
Perspectives on CS	71	0	Too few non-passers to run test
Programming 2	50	21	OPTIMISM, SELFEST (both higher for the passers)
Usability	64	7	No factors differ
Algorithms & data structures 1	45	26	OPTIMISM, SELFEST (both higher for the passers)
Web technology	11	60	No factors differ
Computer architecture	47	24	SELFEST (higher for the passers)
Algorithms & data structures 2	40	31	OPTIMISM, SELFEST (both higher for the passers)
Programming languages	5	66	Too few passers to run test
Regularity & automata	44	27	SELFEST (higher for the passers)

We would liked to have looked for differences between the nine students who passed all 12 courses and the rest. Unfortunately, the number of students who have completed the first year and answered our questionnaire was too low to allow for statistical analysis (the number was five).

Discussion

We expected to find a positive correlation between social well-being and emotional health and how well students perform in introductory computer science courses at university. We did not find a correlation, and that of course is a surprise; particularly because others have found PERFECT, SELFEST and POM to be predictors of success.

Out of curiosity, we conducted a similar study for first-year math students, and our findings were basically the same ($n = 50$). The results of the Pearson correlation tests show that none of the variables correlates with the results of exams (using a threshold of 0.3). The closest r is -0.289 between POM and 'Calculus 2'. The negative sign indicates that students with a more positive affective state perform worse!

In general we have found no linear correlation between the exam score and the social well-being and emotional health. However, in most courses it seems like the students who pass have a higher self-esteem than those who do less well (Table 7). The more positive students seem to do better as well, but it is only significant in three of the courses where a statistical test could be performed. This points toward a positive impact of changes in pedagogy or other aspects aiming at raising students' self-esteem.

The statistical powers of our findings are not impressive—see Table 6. Indeed there might be an effect of the variables on the success of computer science even though we did not confirm it. To be able to be more definitive, we need to have a larger population. This could be achieved by repeating the study in the next semesters and/or by including students who have answered the questionnaire and taken one or more exams later (i.e., in their second year).

Due to the non-conclusive result of our research, we can neither accept nor abandon our hypothesis. It is therefore relevant to identify models that can explain our observations and the lack of confirmation of the hypothesis and previous findings.

Verification of knowledge is traditionally discussed in terms of validity, reliability and generalizability (Silverman, 2001).

- *Validity*—the accuracy of research findings (do the variables measure what we claim).
- *Reliability*—consistency of research findings (will replicated measuring give the same result).
- *Generalizability*—are the results transferable to new contexts.

Our discussion of potential explanation models is structured along these three terms.

Validity

Before the study, we anticipated that the test-instruments were applicable in a Danish context. We focused on the problems of translating the test instruments from English to Danish and expected the construct bias to be negligible (van de Vijver & Hambleton, 1996). Consequently, we focused on the translation of the words to ensure the correct meaning. In retrospect, the test instruments we have used are perhaps less suitable for Danish students due to cultural differences between Denmark and the US. Quite a few of the students responded negatively to the questionnaire and mentioned that they found many questions to be 'odd' and 'out of place'.

Reliability

At least one of our students responded that she found it difficult to indicate the representativity of a statement such as 'I feel I must perform to perfection and if I can't, I rather don't perform at all' without a definition of the intended situation. The statement may be more true in the context of the test person's family, and less true in the context of friends, or vice versa. In either case, the question is ambiguous, and that made some students uncomfortable about the questionnaire as a reliable test instrument.

We have no reason to question the general reliability of the test instruments, but in retrospect, we could have been more specific about the intended situations to consider when answering questions about emotions.

Generalizability

Our findings do not confirm previous findings but that does not allow us to conclude a contradiction with earlier findings. That conclusion is valid only if we can generalize previous findings to our context (or vice versa), which we cannot. A number of circumstances are sufficiently different that a generalization is not immediately possible.

Freshmen at Danish universities are generally a few years older than freshmen at US colleges. The difference in age is likely to imply a difference in maturity and degree of dependency on parents which may be reflected in the way students responded to the questionnaire.

In Denmark, university education is free; furthermore, students receive financial support from the state. In the US the average cost of a private four-year college was 23,940 USD per year (US Department of Education, 2003) and 9828 USD per year in a public four-year college in 2003. Furthermore, parents often contribute some or all of this cost, and support their children financially as well. The resulting stronger dependency between US students and their parents may cause US students to feel stronger pressure from their parents whether the pressure is real or not.

The culture in Denmark is different from the culture in the US. Denmark is a land of homogeneity and equality much more than the US. In Denmark it is considered important that everyone has equal rights and opportunities regardless of social background. Also, it is considered inappropriate if someone stands too much out from the crowd (there is a prevailing who-do-you-think-you-are attitude often caricatured in Danish literature); in the US, this character of personality seems to be encouraged and highly appreciated. Cultural differences make it questionable whether it is possible to generalize US findings to a Danish context (and vice versa); consequently, we cannot conclude a contradiction with earlier findings.

If the above speculations carry some truth, it is necessary to apply a different test instrument more suitable to the Danish context in order to properly test our hypothesis of a positive correlation between social well-being and emotional health and students' performance in introductory computer science courses at university.

Conclusion

To help focus our teaching resources, we have been looking for factors that are predictors of success and which can help improve students' learning premises.

We have studied the influence of emotional and social factors on students' learning outcomes. We have measured the factors in terms of five variables: perfectionism, self-esteem, coping tactics, affective states and optimism.

No correlation was found; the variable that predicts the best is optimism with respect to a course on 'Regularity & automata', but the prediction is very weak: the Pearson correlation coefficient is only 0.342 (0.5 is considered to be the threshold value for correlation). However, the statistical power of the correlations was weak. In most of the courses there were a statistically significant difference between the students who passed the exam and the rest of the students; this corresponds to the findings of related research.

Others have found the variables perfectionism, self-esteem and affective states to be predictors of success. In terms of validity, reliability and generalizability we have identified potential explanations for this seeming contradiction.

Future work

It is a noble goal to identify factors that are predictors of success and which may help improve students' learning premises. Since the statistical power of this study is as limited as described, we need to continue to collect data in order to either verify or reject our hypothesis of a positive impact of social and emotional well-being.

To be able go generalize our findings it would be relevant to repeat the study in different cultural settings—to do a multi-institutional study. In the future, we will seek fellow researchers who are interested in doing so.

Acknowledgement

We would like to thank the students at university of Aarhus who volunteered to participate in the study. Without their help, the study would not have been possible. We would furthermore like to thank the anonymous reviewers of this paper for their thorough comments and suggestions for improvements.

Notes

1. ACT: is formerly known as the American College Test. An American, nation-wide college entrance exam. It assesses high school students' general educational development and their ability to complete college-level work. It is a multiple-choice test that covers four skill areas: English, mathematics, reading and science. The writing test, which is optional, measures skill in planning and writing a short essay (ACT, n.d.). SAT (formerly known as the Scholastic Aptitude Test and Scholastic Assessment Test) is a standardized reasoning test taken by US high school students applying for college. It covers two areas—verbal and mathematics (SAT, n.d.).
2. ECTS: European Credit Transfer and Accumulation System. A full year of study is 60 ECTS points.

References

ACT (n.d.) *ACT assessment.* Available online at: www.act.org/aap/ (accessed 31 August 2006).

Argyle, M. (1994). *The psychology of interpersonal behaviour.* London: Penguin.

Bennedsen, J., & Caspersen, M. (2005). An upcoming study of potential success factors for an introductory model-driven programming course. *Proceedings for the Fifth Koli Calling Conference on Computer Science Education* (pp. 166 – 169). Koli: Finland.

Bennedsen, J., & Caspersen, M. (2006). Abstraction ability as an indicator of success for learning object-oriented programming? *SIGCSE Bulletin (Association for Computing Machinery, Special Interest Group on Computer Science Education)*, *38*(2), 39 – 43.

Bennedsen, J., & Caspersen, M. E. (2005). An investigation of potential success factors for an introductory model-driven programming course, *ICER '05: Proceedings of the 2005 International Workshop on Computing Education Research* (pp. 155 – 163). Seattle, WA: USA.

Bosson, J. K., Swann, W. B. J., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: the blind men and the elephant revisited? *Journal of Personality and Social Psychology*, *79*(4), 631 – 643.

Brooks, J. H., & DuBois, D. L. (1995). Individual and environmental predictors of adjustment during the first year of college. *Journal of College Students Development*, *36*, 347 – 360.

Butcher, D. F., & Muth, W. A. (1985). Predicting performance in an introductory computer science course. *Communications of the ACM*, *28*(3), 263 – 268.

Carver, C. S. (1997). You want to measure coping but your protocol's too long: consider the brief COPE. *International Journal of Behavioral Medicine*, *4*(1), 92 – 100.

Chamillard, A. T. (2006). Using student performance predictions in a computer science curriculum, *ITICSE '06: Proceedings of the 11th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education* (pp. 260 – 264). Bologna: Italy.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* Mahwah, NJ: Lawrence Erlbaum Associates.

Corbetta, P. (2003). *Social research. Theory, methods and techniques.* London: Sage.

Denning, P. J. (2004). The field of programmers myth. *Communications of the ACM*, *47*(7), 15–20.

Dijkstra, E. W. (1969). *Notes on structured programming.* No. EWD 249.

Engel, G., & Roberts, E. (2001). *Computing curricula 2001 computer science, final report.* Available online at: www.acm.org/education/curric_vols/cc2001.pdf (accessed 31 May 2007).

European Union (2006). *ECTS—European credit transfer and accumulation system.* Available online at: http://ec.europa.eu/education/programmes/socrates/ects/index_en.html (accessed 16 August 2007).

Evans, G. E., & Simkin, M. G. (1989). What best predicts computer proficiency? *Communication of the ACM*, *32*(11), 1322–1327.

Exam-scale (n.d.). www.retsinfo.dk/_GETDOCM_/ACCN/B19950051305-REGL (accessed 4 August 2006).

Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, *2*, 15–34.

Gries, D. (1974). What should we teach in an introductory programming course?, *SIGCSE '74: Proceedings of the Fourth SIGCSE Technical Symposium on Computer Science Education* (pp. 81–89).

Hiim, H., & Hippe, E. (2006). *Undervisningsplanlægning for faglærere* [Planning teaching for subject teachers. in Danish]. Copenhagen: Gyldendal.

Lawrence, D. (1996). *Enhancing self-esteem in the classroom.* London: Chapman.

Leafgran, F. A. (1989). Health and wellness program. In M. L. Upcraft & J. N. Gardner (Eds.), *The freshman year experience* (pp. 156–167). San Francisco: Jossey-Bass.

Lee, S., Lee, A. M., & Leung, T. (1998). Cross-cultural validity of the eating disorder inventory: a study of Chinese patients with eating disorders in Hong Kong. *International Journal of Eating Disorders*, *23*(2), 177–188.

Madsen, K. S. (2006, July 31). Politikere: Faldende it-optag skal vendes [Politicians: The falling enrolment to IT studies has to be turned]. *Computer World Denmark*. Available online at: (accessed 31 August 2006).

McCracken, M., Almstrum, V., Diaz, D., Guzdial, M., Hagan, D., Kolikant, Y. B., et al. (2001). A multi-national, multi-institutional study of assessment of programming skills of first-year CS students. *SIGCSE Bulletin (Association for Computing Machinery, Special Interest Group on Computer Science Education)*, *33*(4), 125–180.

McGettrick, A., Boyle, R., Ibbett, R., Lloyd, J., Lovegrove, G., & Mander, K. (2005). Grand challenges in computing: education—a summary. *The Computer Journal*, *48*(1), 42–48.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). *Manual for the profile of mood state.* San Diego: Educational and Industrial Testing Service.

Mouw, J. T., & Khanna, R. K. (1993). Prediction of academic success: a review of the literature and some recommendations. *College Student Journal*, *27*(3), 328–336.

Norem, J. K., & Cantor, N. (1986). Defensive pessimism: harnessing anxiety as motivation. *Journal of personality and social psychology*, *51*(6), 1208–1217.

Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students. Findings and insights from twenty years of research.* San Francisco, CA: Jossey-Bass.

Pritchard, M. E., & Wilson, G. S. (2003). Using emotional and social factors to predict student success. *Journal of College Student Development*, *44*(1), 18–28.

Raad, B. D., & Schouwenburg, H. C. (1996). Personality in learning and education: a review. *European Journal of Personality*, *10*(5), 303–336.

Rice, K. G., & Mirzadeh, S.A.L.I. (2000). Perfectionism, attachment, and adjustment. *Journal of Counselling Psychology*, *47*(2), 238 – 250.

Robins, A., Rountree, J., & Rountree, N. (2003). Learning and teaching programming: a review and discussion. *Journal of Computer Science Education*, *13*(2), 137 – 172.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Rountree, N., Rountree, J., Robins, A., & Hannah, R. (2004). Interacting factors that predict success and failure in a CS1 course, *ITiCSE-WGR '04: Working Group Reports from ITiCSE on Innovation and Technology in Computer Science Education*, Leeds (pp. 101 – 104).

Sanders, R. T. Jr. (1998). Intellectual and psychosocial predictors of success in the college transition: a multiethnic study of freshman students on a predominantly white campus, University of Illinois at Urbana-Champaign. *Dissertation Abstracts International Section B: The Sciences and Engineering*, *58*(10-B), 5655.

Sanna, L. J. (1998). Defensive pessimism and optimism: the bitter-sweet influence of mood on performance and prefactual and counterfactual thinking. *Cognition and Emotion*, *12*, 635 – 665.

SAT (n.d.). *Encyclopædia britannica.* Available online at: www.britannica.com/eb/article-9377802 (accessed 31 August 2006).

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the rosenberg self-esteem scale in 53 nations: exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, *89*(4), 623 – 642.

Shin, Y. H., & Colling, K. B. (2000). Cultural verification and application of the profile of mood states (POMS) with Korean elders. *Western Journal of Nursing Research*, *22*(1), 68 – 83.

Silverman, D. (2001). *Interpreting qualitative data. methods for analysing talk, text and interaction.* London: Sage Publications.

Soloway, E., & Spohrer, J. C. (1989). *Studying the novice programmer.* Hillsdale, NJ: Lawrence Erlbaum.

Soper, D. S. (2007). *A-priori sample size calculator (multiple regression) free statistics calculators (online software)*. Available online at: www.danielsoper.com/statcalc (accessed 20 April 2007).

Sprague, P., & Schahczenski, C. (2002). Abstraction the key to CS1. *Journal of Computing in Small Colleges*, *17*(3), 211 – 218.

Szulecka, T. K., Springett, N. R., & de Pauw, K. W. (1987). General health, psychiatric vulnerability and withdrawal from university in first-year undergraduates. *British Journal of Guidance and Counselling Special Issue: Counselling and health*, *15*, 82 – 91.

Ting, S. R., & Robinson, T. L. (1998). First-year academic success: a prediction combining cognitive and psychosocial variables for Caucasian and African American students. *Journal of College Students Development*, *39*, 599 – 610.

Tucker, A. B. (1996). Strategic directions in computer science education. *ACM Computing Surveys*, *28*(4), 836 – 845.

US Department of Education (2003). *Think college early: average college costs.* Available online at: www.ed.gov/students/prep/college/thinkcollege/early/students/edlite-college-costs.html (accessed 19 April 2007).

van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: some practical guidelines. *European Psychologist*, *1*(2), 89 – 99.

Van Heyningen, J. J. (1997). Academic achievement in college students: what factors predict success? *Dissertation Abstracts International Section A: Humanities & Social Sciences*, *56*(6-A), p. 2076.

Ventura, P. (2005). Identifying predictors of success for an objects-first CS1. *Computer Science Education*, *15*(3), 223 – 243.

Weeks, A., Swerissen, H., & Belfrage, J. (2007). Issues, challenges, and solutions in translating study instruments. *Evaluation Review*, *31*(2), 153 – 165.

Wiggins, N., & Blackburn, M. (1969). Prediction of first-year graduate success in psychology: peer ratings. *The Journal of Educational Research*, *68*(2), 81–85.

Wolfe, R. N., & Johnson, S. D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement*, *55*(2), 177–185.

Yeun, E. J., & Shin-Park, K. K. (2006). Verification of the profile of mood states-brief: cross-cultural analysis. *Journal of Clinical Psychology*, *62*(9), 1173–1180.