An Assessment of Computer Awareness and Literacy among Entry-Level University of Colombo Undergraduates: A Case Study

Thilaksha H. Tharanganie, W. N. Wickremasinghe and G. P. Lakraj

***Abstract* — As the demand for the computer literate is increasing at a rapid pace, possessing of computer skills is an important asset for a university student. Thus having a good computer knowledge improves the quality of their study programs. This paper discusses a case where information was collected through a survey to assess the computer knowledge of entering freshmen in five Faculties (Science, Arts, Management & Finance, Law and Medicine) of the University of Colombo, Sri Lanka. A survey was conducted among 300 new entrants of the above Faculties. A descriptive analysis was used to identify the patterns of computer usage. It is found that from the entry-level University of Colombo undergraduates, majority of students have used a computer (93%) and/or Internet (60%). Moreover, 60% are computer aware while only 47% are computer literate. It must be noted that males in general outperformed females in computer awareness, computer literacy and Internet usage. Since Chi-Square test confirmed that the two variables, computer awareness and computer literacy are associated with each other, rather than considering these two variables separately, considering them jointly is or was??? expected to yield better results. Hence, the two variables are combined into one, with 4 levels and a generalized logit model isor was??? fitted to this nominal multi-category variable to find the factors affecting computer awareness and/or computer literacy. The model suggests that the new variable is dependent on the factors: usage of Internet, monthly family income level, methods of obtaining computer knowledge, and locations of using a computer.**

***Index Terms* — Computer awareness, Computer literacy, Nominal categorical data, Generalized Logit model**

I. Introduction

Technology plays an important role in accelerating economic growth and promoting development. Perhaps no other single technological innovation during the second half of the 21st century that has touched so many lives, than the computer [1].

Manuscript received May 01, 2010. Recommended by Dr. K.P. Hewagamage on May 18, 2011.

Thilaksha H. Tharanganie is with the Department of Statistics, University of Colombo, 35, Reid Avenue, Colombo-7, Sri Lanka. (e-mail: thilakshat@yahoo.com).

W. N. Wickremasinghe and G. P. Lakraj are also with the Department of Statistics, University of Colombo, 35, Reid Avenue, Colombo-7, Sri Lanka. (e-mail: wnw@stat.cmb.ac.lk, pemantha@stat.cmb.ac.lk).

With the increased use of information and communication technologies in education, students entering university need basic computer skills. As students come from different socio-economic backgrounds, they have different learning experiences, capabilities, and needs. It is rarely the case that computer skills of university freshmen are at the same level.

Computer literacy is a mixture of awareness (eg: awareness of the computer's importance), knowledge (eg: knowledge of what computers are and how they work), and interaction (ability to interact with computers). This view is embraced by [2], [3], and [4]. The perspective of computer literacy in [5], involves conceptual knowledge related to basic terminology (including social, ethical, legal, and global issues) and skills necessary to perform tasks in word processing, database, spreadsheets, presentation, graphics, and basic operating system functions.

A search for finding out the factors affecting computer awareness and computer literacy by modeling responses with suitable models did not produce favorable results. However, the review revealed a small number of studies of similar nature. One study was conducted by the Temple University [6] with 259 entry level students to determine their computer literacy at the beginning of the 1997-98 academic years. The study used a questionnaire and revealed the following:

At least 60% of the entering students had access to computers. Approximately 60% of the students had used e-mail, online information services, or the World Wide Web. Students used word processing most frequently and database management systems software least often. In [7], results from a computer concepts assessment given to students enrolled in a computer literacy course at Midwestern Regional University were reported. Slightly over 75% of these students scored more than the minimum college entrance acceptable score on word processing and 63.55% for presentation skills but only 40% for database. Only 6.13% of these students had college entrance scores that exceeded the minimum considered to be acceptable for all components of the test, which also included a wide range of additional topics (networking and the Internet, social and ethical issues, presentation, graphics, operating systems/hardware, word processing, database, and spreadsheets). Besides, all students had a vague idea of

selected computer terms, with some variation by discipline. Another study [8], was a written questionnaire for incoming medical students at the School of Medicine Virginia Commonwealth University from 1991 to 2000. The survey's purpose was to learn the students' levels of knowledge, skill, and experience with computer technology to guide instructional services and facilities. The questionnaire was administered during incoming medical students' orientation or mailed to students' homes after matriculation. The average survey response rate was 81% from an average of 177 students. Six major changes were introduced based on information collected from the surveys and advances in technology: distribute CD-ROMs containing required computer-based instructional programs, delivery of evaluation instruments via the Internet, modification of the lab to PC-based environment, development of an electronic curriculum website, development of computerized examinations to prepare them for the computerized national board examinations, and initiation of a Personal Digital Assistant (PDA) project.

This paper is based on a survey to assess the computer knowledge of entering freshmen in five Faculties of the University of Colombo, in order to determine if incoming students possess the basic computer knowledge and skills to begin studies effectively. The survey also identifies the factors affecting new entrants' computer awareness and computer literacy. Thus, this study provides the computer usage of freshmen in different aspects but does not focus on the knowledge of the undergraduates studying from second year to fourth year. The survey was carried out within freshmen's first three months of the entry to the university in the year 2009 and a sample of 300 was selected using Stratified random sampling. They were interviewed using face-to-face interviewing mode using a questionnaire containing questions of the type single-choice, multiple-choice, 1-3 ranking (1-highest to 3-lowest), and 5-point Likert-scale (1-strongly disagree to 5-strongly agree). Initially a descriptive analysis was done followed by an advanced analysis that resulted in fitting a generalized logit model.

II. Methodology

A. Sampling Procedure

Sampling procedure for our study contains two steps in selecting the sample size. Firstly using the stratified random sampling [9][10], different strata were identified and secondly using proportional allocation method [9][10], the sample size for each stratum was calculated. In this study, five Faculties were considered as five strata and then the sample size within each stratum was divided considering gender frequency. Then, total sample size 300 was divided among the ten strata such that the sample size is proportional to the population size. A benefit of stratified sampling is that estimates will be more precise since each stratum is more homogeneous (less variable) than the population as a whole. Table I displays the sample size allocated for each stratum.

TABLE I
SAMPLE SIZE FOR EACH STRATUM USING PROPORTIONAL ALLOCATION METHOD

Strata	Gender	Population size	Sample size	Percentage of total sample size 300
Science	Male	253	39	13.00
	Female	221	35	11.67
Arts	Male	139	22	7.33
	Female	508	78	26.00
Law	Male	34	5	1.67
	Female	168	26	8.67
Management & Finance	Male	166	26	8.67
	Female	237	37	12.33
Medicine	Male	106	17	5.67
	Female	98	15	5.00
Total	Male	697	109	36.33
	Female	1230	191	63.67

In this sampling method, the simple random sampling [9][10] was performed independently within each stratum. According to the Table I, a simple random sample of 39 Male students had to be taken from Science Faculty for the data collection. Likewise, 10 simple random samples of different sizes were required. Since we use the proportional allocation method, the sample sizes are calculated proportionately to the population size, different sample sizes were obtained. If the strata are about the same size then it will be more convenient to take the same sample size in each stratum.

In this study, a particular procedure was used to make certain, the samples are random and without bias. At first, contact details of male and female students of the five Faculties were obtained and then the specified number of students from each stratum was selected using a random number table, and only those selected, were contacted later for an interview. If a particular student was unwilling to respond, another student was contacted. Interviews were through direct face-to-face interaction to minimize possible questions and confusions that may arise during the process of answering the questionnaire. However, the respondent filled the questionnaire while the interviewer helped them with clarifications when necessary. At the time of collecting questionnaires, it was ensured that all required fields were answered.

B. Computer Awareness

A few definitions are available in literature for the term 'computer awareness' worldwide [1] [11]. In [1], if a person has heard of at least one of the uses of a computer (eg: playing games to complicated aeronautic applications), then he/she is considered a person with computer awareness. Another study [11], has used five pointers: short history of computers, short history of Internet, ways computers are used in the society, occupations related to computer use, and computer ethics, in order to measure the awareness.

In this study, five indicators were created with the help of the above two studies. Five pointers of the study [11] were

adjusted in a manner so that the respondents can understand the indicators easily. These indicators were created in order to go with the knowledge of the new entrants of the University of Colombo. Hence, a person is considered "computer aware" if he/she possesses all of the following five indicators:

- Knowledge about the fundamentals of computers (i.e. Hardware and software computer systems, Computer generations etc.).
- Knowledge about the fundamentals of Internet (i.e. What is Internet, What are the services offered by the Internet? etc.).
- Knowledge related to computer concepts such as social, ethical and legal issues.
- Knowledge about at least three ways that computers are used in society.
- Knowledge about at least three occupations related to computer usage.

C. Computer Literacy

There is no precise consensus on how to define computer literacy [1] [11] [12] [13]. The term can mean different things to different people. Since the respondents of this study are university freshmen, the technical definitions may not be appropriate. After a careful search in literature, it was understood that the computer literacy is defined with three types of skills: basic, intermediate, and advanced. Basic and intermediate skills are being able to use basic operating system functions, word processor, spreadsheets, presentation graphics, databases, Internet, and e-mail. Advanced skills include programming, fixing software conflicts, and repairing computer hardware etc.

Since advanced skills demand technical knowledge beyond the level of a university freshman, only the basic and intermediate skills were considered in this study. Hence, a person was considered "computer literate" if he/she possesses all the following six skills:

- Skills in basic hardware and basic operating system functions – Identifying computer parts, powering up and powering down the computer, open/save files, recognize different file types
- Skills in word processing – Create/save/print documents, Insert tables/charts/ labels/symbols, Format page layout (margins, page numbers, page borders)
- Skills in spreadsheets – Create/save/print spreadsheets, Insert tables/charts, Insert functions/formulas
- Skills in presentation graphics – Create/save/print slide shows, Insert new slide/layout/tables/charts, Create animations
- Skills in databases – Design basic databases with queries and reports/forms
- Skills in Internet & e-mail – Surfing the Internet and sending e-mail messages.

It is understood that in measuring the ability to achieve several functions, an assessment is the best option. However due to lack of facilities to conduct an assessment for each and every respondent, (for instance, the field work of the main survey was conducted by only one interviewer in such a way to minimize the interviewer bias and other errors; limited time frame existed to conduct the survey as the assessment is infeasible since it is a time consuming tool; and respondents' unwillingness to grant more time even at the completion of the questionnaire may affect the possibility of performing an assessment successfully), they were interviewed through direct face-to-face interaction in order to minimize the possible questions and confusions during the cause of answering the questionnaire, thus to reduce the gap between the results through an assessment and a questionnaire.

D. Chi-Square Test

The Chi-Square test [14] [15] provides a method for testing the association between two categorical variables in a two-way table. Categorical variables [16] are the ones having two or more categories. In this study, the Chi-Square test was used to test the association between two categorical variables computer awareness and computer literacy. The null hypothesis H_0 assumes that there is no association between the variables (in other words, one variable does not influence to the other variable), while the alternative hypothesis H_1 claims that some association does exist. The Chi-Square test statistic is computed as,

$$X^2 = \sum \frac{(\text{observed frequency} - \text{Expected frequency})^2}{\text{Expected frequency}} \quad (1)$$

where

$$\text{Expected frequency} = \frac{\text{Row total x Column total}}{\text{Grand total}} \quad (2)$$

Chi-Square test is engaged with (r-1)(c-1) degrees of freedom in a two-way table, where r represents the number of categories of the row variable and c represents the number of categories of the column variable in the two-way table. Degree of freedom [17] is the number of values in the final calculation of an equation that are free to vary. When degree of freedom is one, an adjustment known as Yates Continuity Correction [18] must be employed. In this correction, a value of 0.5 is subtracted from the absolute value (irrespective of algebraic sign) of the numerator contribution of each cell. Then the basic Chi-Square computational formula becomes:

$$X^2 = \sum \frac{(\,|\text{observed frequency} - \text{Expected frequency}| - 0.5)^2}{\text{Expected frequency}} \quad (3)$$

An example of a Chi-Square test for a two-way table is given below with the objective of studying the association between smoking habit and heart attack.

Since degrees of freedom = (*r*-1)(*c*-1) = (2-1)(2-1) = 1, equation (3) is used, and $X^2 = (76 - 67.99 - 0.5)^2 / 67.99 + \ldots + (27 - 18.99 - 0.5)^2 / 18.99 = 7.97$.

TABLE II
AN EXAMPLE TO STUDY ASSOCIATION BETWEEN SMOKING AND HEART ATTACK (EXPECTED FREQUENCIES WITHIN BRACKETS)

		Smoking habit		Row total
		Smoker	Non - smoker	
Heart attack	Yes	76 (67.99)	35 (43.01)	111
	No	22 (30.01)	27 (18.99)	49
Column total		98	62	160

This X^2 value is compared with $\chi^2_{1,5\%}$ which is taken from Chi-Square tables at 5% significance level. Since $X^2(7.97) > \chi^2_{1,5\%}$ (3.84), we reject H_0 at 5% level and conclude that there is an association between smoking habit and heart attack. In practice, statistical software is used to perform a Chi-Square test, and in this study, SPSS® [19] statistical software was used.

E. Generalized Logit Model

Generalized Logit Model [20] [21] [22] is used when the variable is nominal multi-category [16] (two or more categories but which do not have an intrinsic order). In this study, Chi-Square test resulted that the variables computer awareness and computer literacy are associated with each other. Hence, it would not be possible to consider them separately. Thus, by joining these two variables, a new variable with four categories was constructed such that these categories are not ordered or ranked. Thus, a generalized logit model was fitted to this new variable with the objective of finding the factors affecting it.

Suppose there is a nominal [16] variable with *I* categories. In fitting the generalized logit model, it is needed to take one of the categories as the 'baseline' so that other categories can be compared according to it. In usual practice, the last category is taken as the baseline, as the comparison will be more meaningful. Thus, when the last category (*I*) is the baseline with a factor *x*, the generalized logit model is,

$$log\left(\frac{p_i}{p_I}\right) = \alpha_i + \beta_i^x, \quad i = 1,\ldots, I\text{-}1 \qquad (4)$$

where p_i is the probability of occurrence of the response of interest (conditional probability) of the i^{th} level of factor *x*; α_i is the intercept of the i^{th} level of factor *x*; and β_i is the parameter estimate of the i^{th} level of factor *x*. Model (4) indicates that the factor *x* affects the nominal variable.

After fitting a generalized logit model, the next step will be to compute the conditional probabilities $\{p_1, p_{2,\ldots,} p_I\}$ so that the vital conclusions can be obtained after examining them. If a model contains *I* categories, then it has (*I*-1) logits. Hence, the model (4) consists of (*I*-1) logits $\{log\left(\frac{p_1}{p_I}\right), log\left(\frac{p_2}{p_I}\right), \ldots, log\left(\frac{p_{I-1}}{p_I}\right)\}$ and using β_i parameter estimates, the values of these logits can be calculated. Finally conditional probabilities can be computed since they satisfy the equation $\sum_{i=1}^{I} p_i = 1$.

F. Analysis Of Single-Choice And Multiple-Choice Questions

In a single-choice question, there is only one response. In a multiple-choice question, there are a number of responses. These responses are usually marked by a 'tick (✓)'. An example for a multiple-choice question is as follows.

Example:

Which locations have you used to make use of computers when you enter the university?

(Multiple answers possible)

a. Home ☐ b. Internet Cafe ☐
c. Study Institution ☐ d. School ☐
e. Friends / relatives place ☐
f. Other (specify)…… ☐

In the analysis of single-choice and multiple-choice questions, first the number of responses of each factor is counted and then percentage of each factor is obtained as follows:

TABLE III
RESULTS OF THE LOCATIONS OF USING A COMPUTER

Factor	Total	Percentage
Home	183	183/638% = 29%
Internet Cafe	56	9%
Study Institution	151	24%
School	143	22%
Friends / relatives place	86	13%
Other	19	3%
Total	638	100%

G. Analysis Of Ranked Responses

Instead of simply choosing the responses using a tick in a single-choice or multiple-choice question, in this type of a question the responses are ranked. In this study, rankings are in 1-3 scale with '1' for the highest rank and '3' for the lowest rank. An example is given below:

Example:

What are the three mostly used software packages when you enter the university?

(Please rank the three most factors 1-highest … 3-lowest)

a. Ms Office packages ☐ b. Database Management ☐
c. Computer Graphics ☐ d. Web Designing ☐
e. Other (specify) …… ☐

As the first step of the analysis, frequencies of the three ranks are counted and then multiplied by weights 0.5 or 0.3 or 0.2 such that the highest rank (i.e. rank 1) is multiplied by the highest weight (i.e. 0.5) and so on. In practice, these weights are chosen such that the sum of the weights is equal to one. Then the total score of each factor is calculated.

Finally percentage of each factor is obtained from the total score of all factors.

TABLE IV
RESULTS OF THE SOFTWARE PACKAGES USED

Factor	Score of the first three ranks			Total score	Percentage
	Rank1 score	Rank2 score	Rank3 score		
a	259*0.5 = 129.5	6*0.3 = 1.8	5*0.2 = 1	132.3	132.3/191.8%= 69%
b	1*0.5 = 0.5	41*0.3 = 12.3	19*0.2 = 3.8	16.6	9%
c	8*0.5 = 4	64*0.3 = 19.2	21*0.2 = 4.2	27.4	14%
d	2*0.5 = 1	13*0.3 = 3.9	25*0.2 = 5	9.9	5%
e	4*0.5 = 2	6*0.3 = 1.8	9*0.2 = 1.8	5.6	3%
Total				191.8	100%

III. RESULTS

A. Descriptive Analysis

1) Computer Usage

Majority of students (93%) have used computers when they enter the university. From these respondents, the survey sought to determine the reasons for computer usage. Since frequency of using a computer is a single-choice question and locations of using a computer is a multiple-choice question, they were analyzed according to Section II F. While the other three factors (purposes of using a computer, software packages used and methods of obtaining computer knowledge) use ranked responses, they were analyzed according to Section II G. The results are given in Table V.

TABLE V
RESULTS OF THE COMPUTER USAGE

Frequency of using a computer	Locations of using a computer	Purposes of using a computer	Software packages used	Methods of obtaining computer knowledge
Several times a week (33)	Home (29)	Educational activities (47)	Ms Office packages (69)	Computer courses (37)
Daily (29)	Study institution (24)	Leisure activities (25)	Computer Graphics (14)	Self study (25)
Once a week (22)	School (22)	Surfing internet (12)	Database Management (9)	School (20)
Rarely (11)	Friends/ relatives place (13)	For e-mails (6)	Web Designing (5)	Another person (9)
Once a month (5)	Internet café (9)	Office work (6)	Other (3)	Family members (8)
	Other (3)	Self-employment (3)		Other (1)
		Other (1)		

From few respondents who have not used computers, reasons for not using a computer were obtained. This is a question with ranked responses and was analyzed according to Section II G. According to the results, the majority (35%) of the respondents has indicated that the main reason for not using a computer is not having a computer at home while for 32% of the respondents the main reason is financial constraints. The results of the analysis are given in Table VI.

TABLE VI
RESULTS OF THE COMPUTER NON-USAGE

Reasons for not using a computer
No computer at home (35)
Financial constraints (32)
Computer usage was difficult (17)
Computer usage was not required (8)
Other (5)
Computer usage was not knowledgeable (3)

2) Computer Awareness and Computer Literacy

According to the definitions for the computer awareness and computer literacy used in this study, an attempt was made to find out the percentages of freshmen having computer awareness and computer literacy. As stated in Section II B and Section II C, a person was considered as computer aware if he/she possesses all five indicators of computer awareness and a person was considered as computer literate if he/she possesses all six indicators of computer literacy. Table VII and VIII show the percentages of freshmen having computer awareness and computer literacy. The study shows that 60% of the new entrants are computer aware whereas only 47% are computer literate. It is important to note that males in general surpass females in both computer awareness and computer literacy.

TABLE VII
RESULTS OF COMPUTER AWARENESS

	Total number	Number of computer aware	Percentage of computer aware
Male	109	71	71/109% = 65%
Female	191	108	108/191% = 57%
Total	300	179	179/300% = 60%

TABLE VIII
RESULTS OF COMPUTER LITERACY

	Total number	Number of computer literate	Percentage of computer literate
Male	109	55	55/109% = 50%
Female	191	85	85/191% = 45%
Total	300	140	140/300% = 47%

3) Internet Usage

Out of the 278 freshmen who have used a computer when they enter the university, around 60% are Internet users and they mainly use Internet for educational purposes. In addition, 30% students use the Internet several times a week, while a significant proportion of students (26%) are rare Internet users. The results of the Internet usage are given in two Tables below.

TABLE IX
RESULTS OF THE INTERNET USAGE

	Total number	Number of Internet users	Percentage of Internet users
Male	101	64	64/101% = 63%
Female	177	100	100/177% = 56%
Total	278	164	164/278% = 59%

TABLE X
RESULTS OF THE INTERNET USAGE

Purposes of using Internet	Frequency of using Internet
Education & learning activities (37)	Several times a week (30)
For getting information (25)	Rarely (26)
Leisure activities (22)	Once a week (21)
Communication (14)	Daily (14)
Office work (1)	Once a month (9)
Self employment (1)	
Other (0)	

B. Testing The Association Between Computer Awareness And Computer Literacy

In order to fit suitable models for the two variables computer awareness and computer literacy, it was first needed to find out whether they are associated or not. Chi-Square test [14] [15] is used for testing the association between two or more categorical variables. Here, computer awareness and computer literacy are two categorical variables with two categories each.

Categories	Computer awareness	Computer literacy
Category 1	Computer aware (YES)	Computer literate (YES)
Category 2	Non-computer aware (NO)	Computer illiterate (NO)

Thus in order to find the association between these two variables, Chi-Square test was used. The hypothesis of the test is, H_0: Computer awareness and Computer literacy are not associated vs H_1: Computer awareness and Computer literacy are associated. Since the degrees of freedom is one ($(r\text{-}1)^*(c\text{-}1)$ = (2-1)*(2-1) = 1), the Yates Continuity Correction was used (Section II D). Results of the Chi-Square test obtained from SPSS® [19] statistical software are as follows.

TABLE XI
RESULTS OF THE CHI-SQUARE TEST

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	21.817	1	.000	
Continuity Correction	20.676	1	.000	
Likelihood Ratio	22.176	1	.000	
Fisher's Exact Test				.000
Linear-by-Linear Association	21.738	1	.000	
N of Valid Cases	278			

A Chi-Square value 20.676 with 1 degrees of freedom (p = .000 < 0.05) illustrates that the test is significant (reject H_0) at 5% level and there is significant evidence to confirm that the two variables are associated with each other.

C. Fitting A Generalized Logit Model

The above result (Table XI) suggests that computer awareness and computer literacy have to be considered jointly rather than separately. In order to consider them jointly, a new variable was created as follows:

Computer awareness	Computer literacy	Categories of a new variable
YES	YES	1
YES	NO	2
NO	YES	3
NO	NO	4

Since the four categories of this new variable are not ordered, a generalized logit model was suitable (Refer Section II E). Moreover, the last category (category 4) was taken as the baseline and SAS® [23] statistical software was used to carry out the model selection.

The forward selection procedure [24] was used in finding the best model. This procedure starts with the null model (intercept term only) and factors that contribute to the new variable are added one at a time and the factor which gives minimum p value is selected. These factors were identified from the questionnaire (Refer Appendix) and they are Gender, District, Family member who does IT related job, Monthly family income, Usage of resources, Purposes of using a computer, Frequency of using a computer, Locations of using a computer, Software packages used, Methods of obtaining computer knowledge, Usage of Internet, Uses of Internet, and Frequency of using Internet. Then, the rest of the factors were added to the selected model (with the most significant factor) and the next most significant factor was selected. This process continues until none of the factors are significant. The final best model was:

$$log\left(\frac{p_i}{p_4}\right) = \alpha + \beta_{ij}^{uin} + \beta_{ik}^{inc} + \beta_{il}^{met} + \beta_{im}^{loc} \quad (5)$$

where

uin = usage of Internet, inc = monthly family income level, met = methods of obtaining computer knowledge, and loc = locations of using a computer;

i = 1, 2, 3;

j = 1 (Internet user), 2 (non-Internet user);

k = 1 (< Rs.15000), 2 (Rs.15000 - Rs.30000), 3 (Rs.30000 - Rs.50000), and 4 (> Rs.50000);

l = 1 (computer courses followed), 2 (school), 3 (self study, family members, another person, other); and

m = 1 (one location), 2 (two locations), 3 (three locations), and 4 (more than three locations).

As there are four levels for the new variable, model (5) consists of three logits (Refer Section II E). In model (5), $\boldsymbol{log}\left(\frac{p_i}{p_4}\right)$; i = 1, 2, 3, are known as logits. Parameter estimates of these logits are provided in Table XII.

After fitting a model, the usual practice is to test adequacy of the model. This aspect of the adequacy of a model is referred to as goodness of fit [24]. The measures used to determine the goodness of fit of the model are Likelihood Ratio Deviance [25] and Pearson Chi-square [26]. The hypothesis of testing the adequacy of the model is, H_0: The model fits the data well vs H_1: The model does not fit the data well. The results of these tests obtained from SAS® [23] show a Chi-Square value of 558.4536 with 741 degrees of freedom (p = 1.000 > 0.05) for Deviance test and a Chi-Square value of 729.6255 with 741 degrees of freedom (p = 0.6101 > 0.05) for Pearson test, suggesting that the model (5) fits the data well at 5% significance level.

TABLE XII
PARAMETER ESTIMATES OF THREE LOGITS OF MODEL (5)

Factor	Factor levels	Parameter estimates of three logits		
		$log\left(\frac{p_1}{p_4}\right)$	$log\left(\frac{p_2}{p_4}\right)$	$log\left(\frac{p_3}{p_4}\right)$
intercept	α	0.3553	0.1482	-1.1730
uin	j = 1	1.1199	0.1859	1.1392
	j = 2	0.0000	0.0000	0.0000
inc	k = 1	-1.3585	-0.7147	-0.7362
	k = 2	-0.8175	-0.1781	-0.5009
	k = 3	0.5624	-0.0752	0.0822
	k = 4	0.0000	0.0000	0.0000
met	l = 1	0.3578	0.1948	0.5426
	l = 2	-0.2170	0.6758	-0.6477
	l = 3	0.0000	0.0000	0.0000
loc	m = 1	-0.7411	-0.0666	-0.7831
	m = 2	-0.0318	0.2647	0.2148
	m = 3	0.5539	0.8020	0.9503
	m = 4	0.0000	0.0000	0.0000

Conditional Probabilities

After parameter estimation, conditional probabilities for each generalized logit model were calculated, and the following results were obtained.

Model 1: When i = 1, the model (5) becomes,

$$log\left(\frac{p_1}{p_4}\right) = 0.3553 + \beta_{1j}^{uin} + \beta_{1k}^{inc} + \beta_{1l}^{met} + \beta_{1m}^{loc}$$

which models the probability of category 1 of new variable (having both computer awareness and computer literacy) relative to the category 4 of new variable (not having both computer awareness and computer literacy)

Model 2: When i = 2, the model (5) becomes,

$$log\left(\frac{p_2}{p_4}\right) = 0.1482 + \beta_{2j}^{uin} + \beta_{2k}^{inc} + \beta_{2l}^{met} + \beta_{2m}^{loc}$$

which models the probability of category 2 of new variable (having only computer awareness) relative to the category 4 of new variable (not having both computer awareness and computer literacy)

Model 3: When i = 3, the model (5) becomes,

$$log\left(\frac{p_3}{p_4}\right) = -1.1730 + \beta_{3j}^{uin} + \beta_{3k}^{inc} + \beta_{3l}^{met} + \beta_{3m}^{loc}$$

which models the probability of category 3 of new variable (having only computer literacy) relative to the category 4 of new variable (not having both computer awareness and computer literacy)

Then a total of 278 sets of conditional probabilities {p_1, p_2, p_3, p_4} have to be calculated for the 278 respondents who have used a computer when they enter the university. In order to do this for different j, k, l, and m values, parameter estimates of twelve terms (β_{1j}^{uin}, β_{1k}^{inc}, β_{1l}^{met}, β_{1m}^{loc}, β_{2j}^{uin}, β_{2k}^{inc}, β_{2l}^{met}, β_{2m}^{loc}, β_{3j}^{uin}, β_{3k}^{inc}, β_{3l}^{met}, $and\ \beta_{3m}^{loc}$) from Table XII have to be substituted in the above three models. Subsequently, conditional probabilities {p_1, p_2, p_3, p_4} are found using the constraint $\sum_{i=1}^{4} p_i$=1.

Using the conditional probabilities, conclusions for respondents of each category of new variable can be derived. In order to describe category 1, the records which have p_1 as the highest conditional probability are taken from the total of 278 records. Then these highest p_1 conditional probabilities have to be arranged in descending order. Some of them are listed in Table XIII. Altogether 163 records were found with highest conditional probability as p_1. Then, these 163 records were examined to understand about the type of respondents in category 1, and it was found that these records have higher chance of having both computer awareness and computer literacy. Moreover, it is clear that the respondents who are Internet users having higher (Rs.30,000 or greater) monthly family income level and who use three or more locations for computing are more likely to be both computer aware and computer literate. Further, they are likely to obtain computer knowledge from several sources such as computer courses, self study, family members, another person and school.

There were 49 records in which the highest conditional probability is p_2 (category 2 - only having computer awareness). From these records, it can be said that the respondents who are mostly non-Internet users having monthly family income level less than Rs.30,000 and use few (less than 3) locations to use a computer are more likely to be in the computer awareness category only. Additionally, these students obtain their computer knowledge from school and/or by following computer courses.

For category 4 (not having both computer awareness and computer literacy), 66 records were found. From these records, it is evident that the respondents who are mostly non-Internet users having monthly family income level less than Rs.15,000 and use only one location to utilize a computer, seem to be the ones with the highest probability

of not having both computer awareness and computer literacy. Moreover, it appears that their main location of obtaining computer knowledge is school.

It is noted from the results that there was no highest conditional probability found for category 3 (having only computer literacy). This result reveals that there is less probability that a person is only computer literate without being aware of computers.

TABLE XIII
SOME ESTIMATED CONDITIONAL PROBABILITIES FOR CATEGORY 1 OF NEW VARIABLE

	uin	inc	met	loc	p_1	p_2	p_3	p_4
1	1	3	3	3	0.86219	0.07190	0.17928	0.05944
2	1	3	3	4	0.86219	0.01938	0.07732	0.09716
3	1	3	1	1	0.86219	0.17270	0.09679	0.11727
4	1	1	3	2	0.85911	0.15672	0.19829	0.24578
5	1	1	3	4	0.85911	0.04849	0.11967	0.26943
6	1	2	3	4	0.82931	0.03940	0.13140	0.37431
7	1	3	3	3	0.82931	0.07190	0.17928	0.05944
8	1	3	3	4	0.82931	0.01938	0.07732	0.09716
9	1	4	3	1	0.82931	0.11008	0.11905	0.07642
10	1	3	3	1	0.82931	0.09728	0.10216	0.19170
11	1	1	3	4	0.82226	0.04849	0.11967	0.26943
12	1	3	3	4	0.80613	0.01938	0.07732	0.09716
13	1	1	3	4	0.80613	0.04849	0.11967	0.26943
14	1	3	1	4	0.80613	0.03515	0.07483	0.06071
15	1	2	1	3	0.80613	0.24021	0.26728	0.12971
16	1	4	3	1	0.71863	0.11008	0.11905	0.07642
17	1	2	3	3	0.71863	0.13672	0.28505	0.21426
18	1	2	1	2	0.71863	0.23388	0.21345	0.21614
19	1	2	1	3	0.71863	0.24021	0.26728	0.12971
20	1	3	1	4	0.71863	0.03515	0.07483	0.06071
.	.	.	.	.	.	.	.	.

IV. DISCUSSION

This research was carried out with the objective of assessing the computer background knowledge of the University of Colombo freshmen representing five Faculties in the year 2009. Preliminary analysis shows that male students entering the university have higher computer knowledge than their female counterparts. Moreover, a majority of the students have used a computer (93%) and/or Internet (60%) when they enter the university. For those who have not used a computer, not having a computer at home (35%) is the main reason closely followed by financial constraints (32%). About half of the students have used computers mainly for educational and learning activities (47%), and Microsoft Office packages (69%) are the most common in use. A majority has obtained computer knowledge from computer courses (37%), while home (29%) is the most common location of use.

The survey sought to determine the computer and Internet usage of new entrants of the University of Colombo. According to the findings of the study, 60% are computer aware while only 47% are computer literate. It must be noted that males in general do better than females in both computer awareness and computer literacy. From computer users, only 59% are Internet users whereas for them, educational and learning activities (37%) are the main purpose of surfing Internet and most of them use Internet several times a week (30%).

A Chi-Square test was used to identify whether there is an association between computer awareness and computer literacy in order to model the effect of the explanatory variables separately on the two binary response variables computer awareness and computer literacy using two logistic models or jointly on the two responses using a generalized logit model. The Chi-Square analysis proved that the two response variables are associated with each other, hence a generalized logit model is fitted for the new variable by combining the levels of the two binary responses. The generalized logit model for the combined responses suggests that the combined response is dependent on the factors usage of Internet, monthly family income level, methods of obtaining computer knowledge, and locations of using a computer.

From the research findings, it was revealed that University of Colombo freshmen who are likely to be both computer aware and computer literate possess several characteristics. These respondents are Internet users, and their monthly family income level is high. Further they use more locations for using a computer and they obtain computer knowledge from several sources such as computer courses, self study, family members, another person and school. In contrast, for the respondents who are likely to be both non-computer aware and computer illiterate, it is the other way round; i.e. most of them are non-Internet users from families having low monthly income. In addition, they choose only one location for using a computer and obtain computer knowledge from few sources such as school and/or computer courses.

The analysis further proved that new entrants of the University of Colombo who are likely to be only computer aware, hold following features. They are mostly non-Internet users from families having medium level of monthly income. Besides, they obtain computer knowledge from few types of sources such as school and/or computer courses and use one or two locations for using a computer. A significant finding was that there are no highest conditional probabilities found for respondents who are likely to be only computer literate. Since these values are probabilities, even though the highest conditional probabilities were not found for this category in this study, one cannot conclude that there is no possibility that a person in general to be only computer literate without being aware of computers.

In conclusion, the findings provide evidence that the computer modules at University of Colombo should

concentrate more on improving the computer literacy skill base of students, especially female learners. However, both groups would benefit from further instruction and practical experience in this subject matter. In order to better prepare for university computer modules, the administrative bodies of the university should consider offering practical computer sessions and teach students helpful tips and shortcuts for better computer fluency. Further, administrative bodies can compare these results with results from future classes. Another interesting idea would be to repeat the same survey at the conclusion of the course and compare the pre- and post- results.

ACKNOWLEDGEMENT

The authors wish to sincerely thank all those who supported and participated in this survey.

REFERENCES

[1] Department of Census and Statistics – Sri Lanka. (2008). *Household Computer Literacy Survey of Sri Lanka: 2006/07*. Accessed on December 2008 from http://www.statistics.gov.lk/recent%20publications.htm

[2] Hess C. A. (1994). Computer literacy: An evolving concept. *School Science and Mathematics*, 94(4): 208-214.

[3] Brock F. J. and Thomsen W. E. (1992). The effects of demographics on computer literacy of university freshmen. *Journal of Research on Computing in Education*, 24(4): 563-570.

[4] Capron H. L. and Johnson J. A. (2004). *Computers: Tools for an information age complete* (8th ed.) Prentice Hall, Inc., New Jersey.

[5] Hindi N. M., Miller D. and Wenger J. (2002). Computer literacy: Implications for teaching a college-level course. *Journal of Information Systems Education*, 13(2): 143-152.

[6] Patrikas E. O. and Newton R. A. (1999). Computer literacy. *Technological Horizons in Education Journal*, 27(5).

[7] Meier R. (2001). CIS 101 Concepts Test at Fort Hays State University. *Kansas Core Outcomes Project*, 10-18.

[8] Schlesinger J. B. and Hampton C. L. (2002). Using a decade of data on medical student computer literacy for strategic planning. *Journal of the Medical Library Association*, 90(2): 202-209.

[9] Barnett V. (1974). *Elements of Sampling Theory*. (3rd ed.) Hodder & Stoughton, London.

[10] Cochran W. G. (1977). *Sampling Technques.* (3rd ed.) John Wiley & Sons Inc., New Jersey.

[11] *K-5 Technology Integration Scope and Sequence*. Accessed on January 2009 from http://www.hhh.k12.ny.us/uploaded/PDFs/Technology/K5_Scope_Seq.pdf

[12] Hindi N. M., Miller D, and Wenger J. (2002). Computer Literacy: Implications for Teaching a College-Level Course. *Journal of Information Systems Education*, 13(2): 143-151

[13] *Computer literacy*. Accessed on December 2008 from http://en.wikipedia.org./wiki/Computer_literacy

[14] *Two-Way Tables and the Chi-Square Test*. Accessed on March 2009 from http://www.stat.yale.edu/Courses/1997-98/101/chisq.htm

[15] *Chi-Square Test for Association using SPSS*. Accessed on March 2009 from http://statistics.laerd.com/spss-tutorials/chi-square-test-for-association-using-spss-statistics.php

[16] *Types of Variable*. Accessed on March 2009 from http://statistics.laerd.com/statistical-guides/types-of-variable.php

[17] *Degrees of freedom (statistics)*. Accessed on March 2009 from http://en.wikipedia.org/wiki/Degrees_of_freedom_%28statistics%29

[18] *MODULE S7 – CHI SQUARE*. Accessed on March 2009 from http://www.okstate.edu/ag/agedcm4h/academic/aged5980a/5980/newpage28.htm

[19] Everitt B. S., and Landau S. (2004). *A Handbook of Statitical Analyses using SPSS*. (1st ed.) Chapman & Hall, CRC Press LLC.

[20] SAS Publishing (2001). *SAS/STAT® Software: Changes and Enhancements, Release 8.2*. (1st ed.) pp.117-128. SAS Institute Inc., North Carolina.

[21] Agresti A. (2002). *Categorical Data Analysis.* (2nd ed.) John Wiley & Sons Inc., New Jersey.

[22] Gerken J. (1991). Generalized logit model. *Journal of Transportation Research Part B: Methodological*, 25(2-3): 75-88.

[23] Geoff D. and Everitt B. S. (2002). *A Handbook of Statistical Analysis using SAS*. (2nd ed.) Chapman & Hall, London.

[24] Collett D. (2003). *Modelling Binary Data*. (2nd ed.) Chapman & Hall, London.

[25] *Likelihood Ratio Tests*. Accessed on February 2009 from http://warnercnr.colostate.edu/class_info/fw663/LikelihoodRatioTests.PDF

[26] *1.3.5.15. Chi-Square Goodness-of-Fit Test*. Accessed on March 2009 from http://itl.nist.gov/div898/handbook/eda/section3/eda35f.htm

APPENDIX

Serial Number	

Questionnaire

Please tick (✓) or rank the appropriate boxes as required and follow the instructions carefully.

Section 1 – Personal Details

1.1 Faculty : …….

1.2 Stream : …….

1.3 Gender : …….

1.4 District of the school you attended during the G.C.E.(A/L) : …….

1.5 District you live when you enter the university: ……..

1.6 Do you have any family member who does IT (Information Technology) related job when you enter the university?

Yes ☐
No ☐

1.7 Monthly income of your family when you enter the university (Rs.):

a. Less than 15,000 ☐
b. 15,000 – 30,000 ☐
c. 30,000 – 50,000 ☐
d. Above 50,000 ☐

Section 2

2.1 Which of the following resources have you used when you enter the university? **(Multiple answers possible)**

a. Radio ☐ b. Television ☐
c. Desktop computer ☐ d. Laptop computer ☐
e. CD ☐ f. Printer ☐
g. Scanner ☐ h. Mobile phone ☐

2.2 Which of the followings have you known about computers when you enter the university? **(Multiple answers possible)**

a. Knowledge about the fundamentals of computer (i.e. Hardware and software computer systems, Computer generations etc.) ☐
b. Knowledge about the fundamentals of Internet (i.e. What is Internet, What are the services offered by Internet? etc.) ☐
c. Knowledge related to computer concepts (e.g. social, ethical and legal issues) ☐
d. At least three ways that the computers are used in society ☐
e. At least three occupations related to computer use ☐

Section 3 – Usage of a computer

3.1 Have you used a computer when you enter the university?

Yes ☐ No ☐

If the answer is Yes, skip to question 3.3. Otherwise, move to the next question.

3.2 What are the three main reasons for not using a computer when you enter the university? **(Please rank the three most factors 1-highest … 3-lowest)**

a. Computer usage was not required ☐
b. Computer usage was not knowledgeable ☐
c. Computer usage was difficult ☐
d. Had no computer at home ☐
e. Financial constraints ☐
f. Other (specify)…… ☐

Now, skip to Section 5.

3.3 Which of the following skills have you possessed when you enter the university? **(Multiple answers possible)**

a. Skills in basic hardware and basic operating system functions (Identifying computer parts, Powering up and powering down the computer, Open/save files, Recognize different file types) ☐
b. Skills in word processing (Create/save/print documents, Insert tables/charts/ labels/symbols, Format page layout-margins, page numbers, page borders) ☐
c. Skills in spreadsheets (Create/save/print spreadsheets, Insert tables/charts, Insert functions/formulas) ☐
d. Skills in presentation graphics (Create/save/print slide shows, Insert new slide /layout/ tables/charts, Create animations) ☐
e. Skills in databases (Design basic databases with queries and reports/forms ☐
f. Skills in Internet & e-mail (Surfing the Internet and sending e-mail messages) ☐

3.4 What are the three main purposes of using a computer when you enter the university? **(Please rank the three most factors 1-highest … 3-lowest)**

a. Education & learning activities ☐
b. Leisure activities ☐
c. Surfing internet ☐
d. For e-mails ☐
e. Office work ☐
f. Self employment ☐
g. Other (specify) …….. ☐

3.5 How often do you use a computer when you enter the university?

a. Daily ☐
b. Several times a week ☐
c. Once a week ☐
d. Once a month ☐
e. Rarely ☐

3.6 Which locations have you used to make use of computers when you enter the university? **(Multiple answers possible)**

a. Home ☐
b. Internet Cafe ☐
c. Study Institution ☐
d. School ☐
e. Friends / relatives place ☐
f. Other (specify) …… ☐

3.7 What are the three mostly used software packages when you enter the university? **(Please rank the three most factors 1-highest … 3-lowest)**

a. Ms Office packages ☐
b. Database Management ☐
c. Computer Graphics ☐
d. Web Designing ☐
e. Other (specify) …… ☐

3.8 What are the three main methods of obtaining computer knowledge when you enter the university? **(Please rank the three most factors 1-highest … 3-lowest)**

a. Computer courses followed ☐
b. School ☐
c. Self study ☐
d. Family members ☐
e. Another person ☐
f. Other (specify) …… ☐

Section 4 – Usage of Internet

4.1 Have you used Internet when you enter the university?

Yes ☐ No ☐

If the answer is Yes, move to the next question. Otherwise, skip to Section 5.

4.2 What is your ability to use Internet when you enter the university?

a. Can use without assistance ☐
b. Can use with assistance ☐

4.3 What are the three main uses of Internet when you enter the university? **(Please rank the three most factors 1-highest … 3-lowest)**

a. Education & learning activities ☐
b. Leisure activities ☐
c. For getting information ☐
d. Communication ☐
e. Office work ☐
f. Self employment ☐
g. Other (specify) …… ☐

4.4 How often do you use Internet when you enter the university?

a. Daily ☐
b. Several times a week ☐
c. Once a week ☐
d. Once a month ☐
e. Rarely ☐

Section 5

5.1 Comment on the following statements.

	Strongly agree	Agree	No opinion	Disagree	Strongly disagree
I think having computer knowledge when enter the university is beneficial.	☐	☐	☐	☐	☐
Having computer knowledge make my life easier.	☐	☐	☐	☐	☐

5.2 Give your comments on any other factors of measuring computer literacy.

……
……
……

Thank You!