Pair Programming: A Teaching and Learning Tool for Defending Student's Mental Energy

Pandarinathan Radhakrishnan[1], Selvadurai Kanmani[2], and Malaiyappan Nandhini[3]
[1]Department of Computer Science and Engineering, Pondicherry Central University, India
[2]Department of Information Technology, Pondicherry Engineering College, India
[3]Department of Computer Science, Pondicherry University, India

Abstract: *Energy is an essential requirement to do any assigned tasks successfully. Mental Energy (ME) is the intellectual power for effective performance of cognitive tasks. This paper discusses the level of ME of the student learning to develop software program using Pair Programming (PP) and compares with the student learning to develop software program using traditional method. Subjective perception of ME was correlated with various elements of program correctness. PP and Solo Programmers (SP) were the two student groups took part in this experimental study. Both groups were asked to do programming assignments as Task 1 (T1) and Task 2 (T2) consecutively without break. Both tasks were given with equal level of software complexity and graded by an automatic tool. The grades were analyzed using non-parametric statistical methods. The results show that, PP group performed both T1 and T2 with high level of program's correctness, scoring almost equal marks in both the tasks. But, the performance of the SP group had more difference between T1 and T2 in program correctness and they scored comparatively less marks in T2 than in T1. The confirmatory analysis by means of questionnaire shown positive correlation over the hypothesis, which implies that, the ME level of PP students remained undiminished until T2, thus proving that PP methodology has more advantage than traditional method SP of learning software program.*

Keywords: *ME, eXtreme Programming (XP), PP, solo programming, automatic grading.*

Received December 26, 2013; accepted October 8, 2014

1. Intoduction

Teaching and learning of software engineering is a major challenge, to integrate theory and practice by some methodology. The rapid technological changes and rapid expansion of available information forces [16] the pedagogic people towards nontraditional teaching and learning methods. When compared to traditional method, pair learning creates stronger, more powerful experiences and accelerates the learning process for students. Being more collaborative and interactive, pair learning has gained immense popularity as a learning tool in academia. In this context, Pair Programming (PP), which is one among the twelve practices of eXtreme Programming (XP) [1], presents an excess benefit, especially in practical learning. Several previous controlled experiments in programming industries claimed that PP is useful and beneficial in numerous facets. PP also contributed several benefits to the program-learning students [6]. Few such benefits are program correctness, higher software quality, reduced time for program development, increased learning efficiency, increased confidence level, increased course completion rate, and so on.

1.1. Pair Programming

PP involves two programmers working collaboratively on one computer, one as a driver who operates the keyboard, concentrates on the lower level details of the task and another as a navigator who observes the driver [1], offering suggestions and corrections on higher level details of the task [11]. When students work, they keep each other on task, and inclined with a greater level of Mental Energy (ME), thereby producing correct programs. In contrary, SP is the traditional method, developing software program individually. If the task is straightforward, solo programming can be more efficient. However, if the assignment is complex, and if struck in-between, he loses his temperament and ME, which hampers his performance.

1.2. Mental Energy

ME is a biological process to do physical work as well as cognitive task effectively [12]. ME is equivalent to general intelligence [2]. Great Intellectual and genius in the field of science as well as other disciplines had an abundance of ME. It has the ability to persist for a longer attention span and a valid predictor of success and achievement [9]. ME is highly correlated with the mental model, which is a cognitive scientific term used

to describe cognitive representations [13], by students, in terms of correct and efficient learning.

The subsequent section describes the background and related work. In section 3 design of the study explained. In section 4 data analysis and the results obtained are discussed. Finally, section 5 concludes the study.

2. Background and Related Works

Several researchers proved that PP increases personality traits of the students [14, 15]. Hannay *et al.* [6] have extensively explored the personality traits, involving the performance of heterogeneous and homogeneous groups. However, different personalities express different kinds of communication, based on information and knowledge, decision making, and problem solving [3, 4] for which they depend heavily on their own ME as shown in Figure 1.

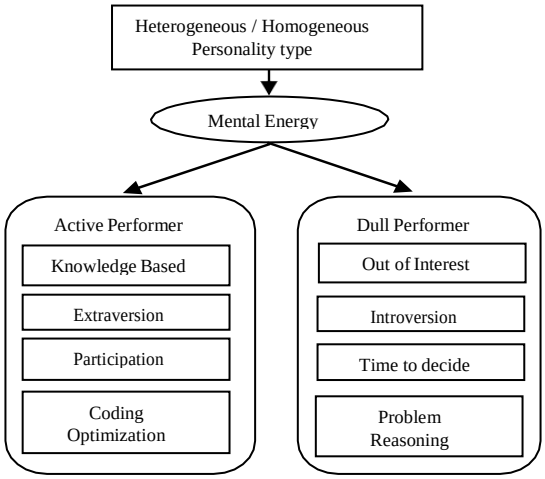

Figure 1. Heterogeneous and homogeneous groups in PP.

Program correctness is one of the most important properties of a software program [5]. It is obvious that the correctness of performance is an intellectual achievement and an outcome of ME [9]. The effective performance by a person on a task becomes ineffectual when the ME is exhausted. An active mind is a vital aspect of attention control and a core ingredient of strong learning. But, student learning the skills of software development in a practical laboratory is sometimes very active and sometimes dull. Interestingly, this active and dull mood refers to the cycles of ME [8]. Many times, the span of this cycle, which incorporates learning and displaying, tends to diminish considerably to due to continuous learning for a long duration, which hampers the learning output.

This might be the evident reason for the authors of XP, the agile methodology for software development [2], suggests forty hours of work in a week as one of its policies. Even though ME is a valid predictor for program correctness as shown in Figure 2, but found no strong indication in the literature that ME correlating student's cognitive performance. The empirical studies on the PP, involving the human factors have not revealed any strategy about the impact of ME while learning software in the laboratories. Hence, there is a need to introduce some innovative and conducive method in the laboratory as an alternative. Such attempts would enable the students to sustain their ME for a long span of time while learning. Therefore the study investigated the correctness of programming output, using the beneficial elements of PP such as: developer's personalities; temperaments on communication; collaboration-viability as definite inducing agents.

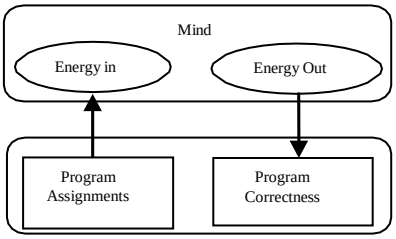

Figure 2. Input and output of the evaluation process.

2.1. ME Measurements

ME of a healthy individual is highest in the morning, falling progressively and significantly over the day [9]. Same is the case of a healthy student attending classes continuously. ME is closely related to the mood of an individual, who performs the task. The mood is influenced by many factors such as environmental conditions, social interactions, and physical activities [11], which yield desired output of tasks [13], and they are accepted as valid variables for measuring ME [10]. ME prevails the outcome of three broad categories for measurement: The amount of workload/task assigned; time to complete the task; and subjective experience. The task is a mental construct and an intervening variable [9] for ME, and it cannot be directly observed. If the task is being measured in an experimental setting, the measurement options are generally wider than those for operational settings [1]. The practical measures could be based on the laboratory studies, and technique acceptable to the students. Measurement of mental task focuses on speed, accuracy, response time and error rates [9], which are the core elements for program correctness and correlating with ME. Performance measures of ME are classified into two major types: primary task and secondary task. Usually, primary task and ME does not correlate, because, under normal situation the performer does not lose his energy during the primary task. Hence, the selection of secondary task becomes mandatory for measuring to determine the correctness.

3. Study Design

Since, this study is designed to evaluate the level of ME of the students', the following research questions framed.

- *Question no 1*: is PP beneficial to students, who learn to do software programming?
- *Question no 2*: is correctness of program is a valid predictor to correlate ME of the students learning software programming?

3.1. Sampling and Pairing

This study was carried out using the first year computer science students of Pondicherry Engineering College. Out of 600 students enrolled in various branches of B.Tech course, 261 first year students who studied computer science as major in HSC were selected. To ensure their equal skill level, the "inter-se merit" list after the normalization method followed by the Centralized Admission Committee (*http://www.centacprof.net.in/ selection.html*), and the marks obtained in the previous test in computer programming lab are considered. Adding both marks, a common rank list was prepared and sorted. Using the sorted array of the rank, 87 pairs were formed for PP by pairing top rank holder with last rank holder iteratively. The remaining 87 students were marked for solo programming as shown in Figure 3.

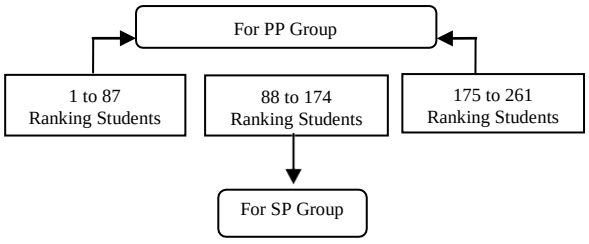

Figure 3. Scheme for selection of students for PP and SP.

3.2. Task Development

As the aim of this experiment is to measure the level of ME, the measuring method should have minimal equipment, to measure the students performance directly. For that, if the task is more than one, the difference could be captured well, only when they are considered as primary and secondary, and carry equal concepts in the demand imposed on the user. Hence, the task equality is confirmed with the program complexity by a pilot study using C programming language. Senior students from B.Tech, Computer Science, took part in the pilot study. To determine the equality of the conceptual complexity, a freeware "Source Monitor-Ver 3.1" was used with the following steps and 6 assignments selected.

Comp= Max. {Conceptual Complexity {T1, T2, T3, …, Tn}}
Input Tasks= {Ti}, if ∀i , Comp= Conceptual Complexity (Ti)

The PP groups were asked to compete any four assignments, two each for task T1 and T2 as they are working in pairs, and SP groups are asked to complete any two assignments, one for each task T1 and T2.

3.3. Automatic Grading Tool

Similar to the selection of students and assignments, attention was taken to grade the program automatically by Computer. Grading the assignments by human is an error-prone task. Hence, a tool called PECA, developed by the authors of this study [7] is used to grade the students program and steps are:

ME ∞ Correctness= program behaviour in {Static and Dynamic} environment
Static = p + r + c + d
where p→ problem definition, r → resource, c → coding analysis and d → debugging
where $\sum_{i=1}^{n} Vi + l + sy + \sum_{i=1}^{n} sol$
where v→ valid input; l→ logical analysis;
sy→ syntax analysis; sol→ feasible solution;
Dynamic=dt/dx+dx/c
where dt/dx→ time t varying with respect to memory x;
dx/dc → memory x varying with respect to code c;

3.4. Variables

The variables used are: independent variable; energy variable; and confirmatory variables. They constitute the constructs for program correctness and considered as components for measuring the ME level.

The programming groups PP and SP are the independent variable. Program Design Score (PDC), Program Coding Marks (PCS), Program Testing Score (PTS) and Total Score (TS), are the energy variables. For confirmatory variable questionnaire is used.

Questionnaire: Since, there were no suitable questions for this experiment available in the literature, five self-administered questions with three points likhert scale were developed for finding the state of ME. Before floating, the students were briefed about the ME and its connection with the task completion.

4. Results and Discussion

First, the mean difference of the total marks, between *T*1 and *T*2 of both PP and SP groups were computed. Second, the mean difference of the total marks of *T*1 in PP and SP and third, the mean difference of the total marks of *T*2 in PP and SP was worked out.

Table 1. Null hypothesis for univariate analysis.

Sl.No	Hypothes
$H1_0$	There is no difference between the scores obtained in T1 and T2 by the PP students
$H2_0$	There is no difference between the scores obtained in T1 and T2 by the SP student
$H3_0$	There is no difference between the scores obtained by PP and SP in T1
$H4_0$	There is no difference between the scores obtained by PP and SP in T2

Keeping in view of the research questions, the hypothesis for the univariate analysis is listed in Tables 1 and 2 for multivariate analysis.

Table 2. Null hypothesis for multivariate analysis.

Sl.No	Hypothesis
$H1_0$	There is no difference between the marks (design, coding and testing) obtained in T1 and T2 by the PP students.
$H2_0$	There is no difference between the marks (design, coding and testing) obtained in T1 and T2 by the SP student.
$H3_0$	There is no difference between the marks (T1 and T2) obtained by the PP and SP students.

4.1. Multivariate Analysis

Multivariate analysis was used to perform trade studies across multiple dimensions while considering the effects of all variables on the responses of interest.

The analytical study was done serially, with the mean difference of energy variables as shown in the conceptual model as shown in Figure 4.

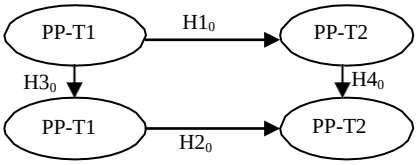

Figure 4. Univariate mean comparisons.

4.2. Explorative Analysis

Explorative analysis consists univariate and multivariate tests. For univariate, "t-Test for the equality of two means" using SYSTAT package applied. Straight and cross validation of the *T*1 and *T*2 performed by both the group were done using TS. For the multivariate analysis, Hoteling's *T* square test, and SPSS tool for output was used.

4.2.1. Univariate Analysis

- *Test 1*: in this test both of the tasks *T*1 and *T*2, performed consecutively by the PP group was analyzed for the null hypothesis *H10* . From the analytical report shown in Table 3, the mean difference of the marks was 3.310 and the p-value was 0.060. Since the p-value was greater than the level of significance value (0.05), there was no evidence to reject the null hypothesis, and show there was no significant difference between the mark obtained in *T*1 and *T*2 by the PP students.
- *Test 2*: in this test, the tasks *T*1 and *T*2 consecutively performed by the SP group were analyzed to test the null hypothesis *H20*. The mean difference of the mark of *T*1 and *T*2 was 6.046 and the p-value was 0.000. Since the p-value was less than the level of significance value (0.05), there was an evidence to reject the null hypothesis. The scores obtained in *T*2 was lesser than *T*1 by the SP student.
- *Test 3*: in this test the total marks of the *T*1 done by PP and SP was analysed to test the null hypothesis *H30*. The analytical report shown in Table 3 reveals that the mean deference and p-values are 14.759 and0.000 respectively. Since p-value was lesser than the level of significance value (0.05), the null hypothesis was rejected.
- *Test 4*: in this test the scores of *T*2 done by PP and SP was considered for analysis. The results shown in Table 3 reveal that the mean deference and p-values are 17.528 and 0.000 respectively. Since, p-value was lesser than the level of significance value (0.05), the null hypothesis *H30* was rejected.

Table 3. Results of univariate analysis.

Univariat Tests	Group	Tasks	N	Mean	SD	MD	95% CB	t-Value	df	p-Value
Test 1	PP	T1	87	71.356	13.693	3.310	-0.186	1.566	172	0.060
		T2	87	68.046	14.189					
Test2	SP	T1	87	56.598	12.110	6.046	3.268	3.600	172	0.000
		T2	87	50.552	9.938					
Test 3	PP, SP	T1	87	71.356	13.693	14.759	11.517	7.530	172	0.000
		T1	87	56.598	13.693					
Test 4	PP, SP	T2	87	68.080	14.180	17.528	14.459	9.442	172	0.000
		T2	87	50.552	9.938					

4.2.2. Multivariate Analysis

- *Test 1*: in test 1, the performance of PP in T1 and T2 was taken into consideration for null hypothesis H10. The marks of the three energy variables design, coding and testing was compared. Since p-value (0.209) shown in Table 4 is greater than the level of significance value (0.05), there is no evidence to reject the null hypothesis.
- *Test 2*: this test is based on the marks of the three energy variables obtained by SP students on *T*1 and *T*2. Since p-value (0.000) shown in Table 4 is lesser than the level of significance value (0.05), the null hypothesis *H*20 was rejected.
- *Test 3*: in this test, all the marks obtained both by PP and SP groups, on *T*1 and *T*2 were used for combined analysis. From the output shown in the Table 4, it was found that the p-value is 0.000, which was the evidence for rejecting the null hypothesis *H*30.

Table 4. Results of multivariate analysis.

Grouping	Tasks	Effect	Value	F	Hypothesis df	Erorr df	P-Value
PP	T1, T2	Hotelling's trace	.027	1.527	3.000	170.000	.209
SP	T1, T2	Hotelling's trace	.554	31.378	3.000	170.000	.000
PP vs. SP	T1, T2	Hotelling's Trace	.590	50.420	2.000	171.000	.000

4.2.3. Confirmatory Analysis

Chi square test is applied to all the five questions, to test the goodness of fit to verify the distribution of observed data with the assumed theoretical distribution. For all the questions, "Type of students" was 87 PP and 87 SP. Answers to the questions were split into three types: Yes; Maybe; and No. Following were the null and alternate hypothesis set to check the association between the two types of students.

- *H0*: there is no association between type of student and answering to the question.
- *H1*: there is an association between type of student and answering to the question.

As shown in Table 5, for the question 1, 85 PP students and 73 SP students answered "Yes"; 2 PP students and 2 SP students answered "No"; and only 12 SP students answered "May be". The Chi-Square and significant values are 12.911 and .002 respectively. For the second question, 85 PP students and 59 SP students answered "Yes"; 2 PP students and 26 SP students answered "May be"; and only 2 SP students answered "NO". The Chi-Square value is 27.266 and the significant value is 000. Since both the significant values are less than 0.05, it was concluded that, there was an association between types of student and answering the question. This showed that, while starting the new program, students were in the high ME level.

Table 5. Results of confirmatory analysis.

Type of students	Yes	May be	No	Total	x	df	Sig.
Were you generally very enthusiastic about starting a new task?							
PP	85	00	02	87	12.911	02	.002
SP	73	12	02	87			
Total	158	12	04	174			
Were you energetic while starting the work?							
PP	85	02	00	87	27.266	02	.000
SP	59	26	02	87			
Total	144	28	02	174			
Were you always "on the go" during programming?							
PP	79	03	05	87	23.550	02	.000
SP	53	25	09	87			
Total	132	28	14	174			
Were you interested, if you have to reform your task more?							
PP	83	02	02	87	22.384	02	.000
SP	59	20	08	87			
Total	142	22	10	174			
Do you agree that break is not required between programming tasks?							
PP	83	02	02	87	10.956	02	.004
SP	69	14	04	87			
Total	152	16	06	174			

For the third question, the Chi-Square value is 23.550 and the significant value is less than 0.05. Therefore the null hypothesis was rejected and concluded that, there was an association between types of student. To this question 79 PP students and 53 SP students answered "Yes"; 3 PP students and 25 SP students answered "May be"; and 5 PP students and 9 SP students answered "NO". Since the question was related to the maintenance of energy, the numbers of SP students expressed "Yes" were considerably less. This showed that, the SP students were losing ME while working on T2.

For the fourth question, the Chi-Square and significant values are 22.384 and .000 respectively. As the significant value is less than 0.05, the null hypothesis was rejected and concluded that, there was an association between types of student and answering to the question. 83 PP students and 59 SP students answered "Yes"; 2 PP students and 20 SP students answered "May be"; and 8 SP students and 2 PP students answered "NO". As the question was to inquire about the readiness of the students to take more tasks, most PP groups have shown interest in taking further task, and most SP groups have shown less interest in taking further tasks, which was added evidence that the PP groups maintain the ME constantly.

For the fifth question the Chi-Square value is 22.384 and the significant value is less than 0.05. Hence, the null hypothesis was rejected and concluded that, there was an association between types of student and answering to the question. To this question 83 PP students and 69 SP students answered "Yes"; 2 PP students and 14 SP students answered "May be"; and 2 PP students and 4 SP students answered "NO". Even though more number of SP students answered that they do not require breaking in between, when compared to PP, more SP students answered May be". This shows that the SP students were on the verge of losing their ME.

4.3. Discussion

To test the level of ME of the students, the energy determinant variables were first fitted to the complete data set, including all categories of developer and tasks. To provide meaningful measures, only correct solutions considered for the analysis. A careful consideration was taken for assigning programs, correcting and providing marks equally. For statistical validity, hypothesis testing using univariate and multivariate analysis were done. In univariate test 1 shows that there was no difference in performance of *T*1 and *T*2. But the Test 2 shows that there was a difference in performance of the SP in *T*1 and *T*2 as shown in Figure 5.

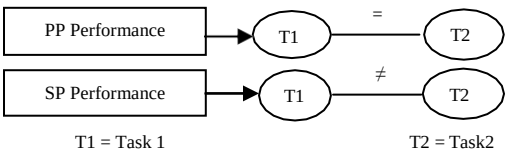

Figure 5. Univariate mean comparisons.

Even though the experimental subjects were selected carefully for their equality in academic skill level, solved equal level of assignments, graded automatically by computer-aided system, the performance level of the SP group in *T*2 reduced considerably than *T*1. But in case of PP group, the performance level remained constant. Evidently, this shows that the SP group was losing their ME level and the PP group maintained constantly.

The univariate tests 5 and 6, the mean difference of T1 and T2 was 14.758 and 17.528 respectively. Even though the difference was expected by virtue of pair performance, the reason for the vast difference is, because of diminishing energy level in SP group.

Multivariate test 3 was conducted to evaluate the whole performance of PP and SP groups in both T1 and T2. The test showed that, the PP group performed

equally in T1 and T2, but the SP groups showed different performance in T1 and T2. The relation between the performance and the ME level is correlated using confirmatory analysis.

Three multivariate tests were conducted using energy variables of design, coding and testing marks. Multivariate test 1 shown that pp group preserved their energy level constantly. In contrary, multivariate test 2 proved that, the SP groups shown difference in performance level of T1 and T2, by upholding the null hypothesis. The performance of T2 by SP group was lesser than T1 which proved that their ME level has come down. But, 95.28% of PP groups conform through their answers, that they hail their ME with the help of PP methodology in software laboratories.

5. Conclusions and Future Work

In this paper, the effects of PP methodology applied in a laboratory, where students learned software programming. It was known for a long time that, learning to program in the computer laboratory by students is difficult. Due to the work pressure, students are likely to lose their ME level considerably. As ME generates effective working capacity of students, this study investigated the level of ME during PP in the laboratory. When students are working in pairs, the interactions, discussions, and the constant support they get from each other play a big role in keeping them in the stable ME throughout the lab session, thus helping them to produce programs correctly, even if the tasks are difficult. On the other hand, the students using traditional method of SP tend to lose their ME gradually, thus lowering their performance in programming in due course.

The research question was addressed by comparing the program correctness and correlating stable ME level of PP students. Further, we need to understand that whether the ME of both the students in the pair is benefitted or not. Whether, this study is viable for both heterogeneous and homogeneous type of students and their mental model. We propose to address these issues in our future study.

References

[1] Beck K. and Andres C., *Extreme Programming Explained: Embrace Change*, Addison Wesley, 2003.

[2] Cain B., "A Review of the Mental Workload Literature," Technical Report Defence Research and Development, 2007.

[3] Cockburn A. and Williams L., *The costs and benefits of pair programming*, Addison-Wesley Longman Publishing Company, 2001.

[4] Ferdinandi P., "Facilitating Communications," *IEEE Software*, vol. 15, no. 5, pp. 92-96, 1998.

[5] Frent M., "Correctness: A Very Important Quality Factor in Programming," *Studia Unvisitey Babes-Bolyai*, Informatica, vol. L, no. 1, pp. 11-20, 2005.

[6] Hannay J., Dyba T., Arisholm E., and Sjøberg D., "The Effectiveness of Pair Programming: A Meta-Analysis," *Science Direct*, vol. 51, no. 7, pp. 1110-1122, 2009.

[7] Kanmani S. and Radhakrishnan P., "A Hybrid Approach to Assess the Students Program Automatically," *Journal of the Indian Institute of Engineers*, vol. 91, pp. 3-9, 2010.

[8] Kulesza T., Stumpf S., Burnett M., and Kwan I., "Tell Me More? The Effect of Mental Model Soundness on Personalizing an Intelligent Agent," *in Proceeding of the Special Interest Group on Computer-Human Interaction Conference on Human Factors in Computing Systems*, Austin, pp. 1-10, 2012.

[9] Laykken D., "Mental Energy," *Journal of Intelligence*, vol. 33, no. 4, pp. 331-335, 2005.

[10] Lieberman H., "Mental energy: Assessing the Cognition Dimension, Nutrition Review," *International Life Sciences Institute*, vol. 64, no. 7, pp. 10-13, 2006.

[11] Maxhuni A., Matic A., Osmani V., and Ibarra O., "Correlation Between Self-reported Mood states and Objectively Measured Social Interactions at Work: A Pilot Study," *in Proceeding of the 5th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*, Dublin, pp. 308-311, 2011.

[12] O'Connor P., "Mental Energy: Assessing the Mood Dimension," *Nutrition Reviews*, vol. 64, no. 7, pp. 7-9, 2006.

[13] Ramli N. and Fauzi S., "The Effects of PP in Programming Language Support," *in Proceeding of International Symposium on Information Technology*, Kuala Lumpur, pp. 1-4, 2008.

[14] Salleh N., Mendes E., Grundy J., and Burch G., "An Empirical Study of the Effects of Conscientiousness in Pair Programming Using the Five-Factor Personality Model," *in Proceeding of the 32nd ACM/IEEE International Conference on Software Engineering*, Cape Town, pp. 577-586, 2010.

[15] Sfetsos P., Stamelos I., Angelis L., and Deligiannis I., "An Experimental Investigation of the Personality Types Impact on Pair Effectiveness in Pair Programming," *Empirical Software Engineering*, vol. 14, no. 2, pp. 187-226, 2009.

[16] Zaied A., "A Framework for Evaluating and Selecting Learning Technologies," *The International Arab Journal of Information Technology*, vol. 4, no. 2, pp. 141-147, 2007.

Pandarinathan Radhakrishnan received his MTech degree in Computer Science and Engineering at Pondicherry Central University. He is currently working in the Department of Computer Science and Engineering, Pondicherry Engineering College.

Malaiyappan Nandhini Assistant Professor in the Department of Computer Science, Pondicherry University, India. She has 17 years of teaching and 10 years of research experience. Her research interests are software engineering, evolutionary algorithms and combinatorial problems optimization.She has published about 60 research papers in journals and conferences.

Selvadurai Kanmani received her BE and ME degree in Computer Science and Engineering from Bharathiar University and PhD from Anna University, Chennai. She has been the faculty of the Department of Computer Science and Engineering, Pondicherry Engineering College since 1992. She has published about 150 papers in international conferences and journals. Her research interests are software engineering and data mining techniques. She is a member of Computer Society of India, ISTE and Institute of engineers, India.