Confusion and Compilation Logs: A Study of Novice Programmer Experiences

Teresita S. Paccarangan
Philippine Science High School
Agham Road, Diliman
Quezon City, Philippines
+632-9291603
tspaccarangan@pshs.edu.ph

ABSTRACT

The study explored the relationship between confusion and compilation behaviour of the novice programmer. Confusion was detected in the context of facial expressions and program writing as captured in video. Videos of student faces and their screens were synchronized, time-sliced and labelled. Percentages of confusion were then correlated with features of the students' compilation behaviours. Our study showed that relationships between video confusion rate and number of errors, video confusion rate and number of compilations, video confusion rate and total time, video confusion rate and EQ, and confusion rate from the compilation logs and the confusion rate from the video observations were not significant. Increasing the affect judgment is recommended by including affect judgment from an expert or who has done affect judgment through videos, and by using smaller time slices for the complete duration of the video.

Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]: Psychology

General Terms

Human Factors

Keywords

Novice programmer, confusion, compilation logs, compilation behavior

1. INTRODUCTION

"Programming Application Development" is the number one skill sought in industry today [25]. However, learning how to program is difficult for many students. Students struggle with their first programming course, and many graduate with little confidence in their programming ability [15] resulting in an alarming drop-out rate [20, 21, 29]. This study is focused on the novice programmer because the experience of the novice programmer is critical. A novice programmer is characterized as a student enrolled in an introductory programming course in college, and is normally labelled as "beginner," "novice", or "novice programmer" [12, 13, 15, 17, 27]. We would like to make sure that the novice programmer's experience will make him want to continue through the advanced programming courses and eventually become a member of the pool of competent software developers much needed in industry.

The goal of the study was to explore the relationship between two things in the experience of the novice programmer: confusion and compilation behaviour. This research attempted to answer the following questions: How can we quantify novice programmer confusion? What is the relationship between confusion and compilation behaviour? The confusion we wanted to picture came from three sources: facial expressions of each student captured on camera, screen capture of the real-time changes in the program that the student was writing, and compilation logs. Side-by-side analysis of the synchronized videos using time-series data with the corresponding compilation data from the logs were used to study the moments or periods of confusion experienced by the novice programmer.

Jadud defined compilation behaviour as the programming behaviour a student engages in while repeatedly editing and compiling their programs [14]. This compilation behaviour was quantified through a specific algorithm based on metrics in the compilation logs [14]. Compilation logs are data sent to the compiler whenever the student compiles his program. In our study, the analysis focused on getting the relationship between confusion rate from the videos, and compilation log data such as number of compilations, number of errors, total time taken in doing the task, average time between compiles, error quotient or EQ, and confusion rate. ELAN, a multimedia annotation tool created by the Max Planck Linguistics Institute, was used for the video analysis [3].

2. RELATED LITERATURE

The group of D'Mello and Graesser were among those who significantly made use of video analysis in studying cognitive-affective behaviour, using an intelligent tutoring system. Their studies stemmed from a progressive research on the field and showed that confusion was common to students engaged in deep learning [4, 20]; confusion was just as persistent as flow and boredom [7]; and that regulated confusion does actually contribute to learning [2, 4, 5, 6, 7, 8, 18, 19, 20, 23]. Retrospective judgment, where the participants identified the specific affect they experienced during specific moments while engaged in learning was one method used to determine affect. In other studies, affect judgment was also done by peer and trained judges while viewing the videos. The most common affective states occurring during deep learning have been identified. Affect judgment was done using this list with their definitions. Confusion was defined as "Noticeable lack of understanding and being unsure about how to proceed."

Studies done by of D'Mello et al. also established the reliability of using the Facial Action Coding System [9] in identifying facial units or muscles of the face that result in facial expressions that occur during deep learning [5, 23]. In these studies [5, 23], the highly animated states of confusion and delight were found to be easily detectable from facial expressions. Confusion was manifested by a lowered brow (AU4), tightening of the eyelids (AU7), and a notable lack of lip corner puller (AU12). Figure 1 below illustrates the faces with action units (AU's) that were identified with the confused affective state [11].

AU4, Brow lowerer   AU7, Lid tightener   Notable lack of AU12, Lip corner puller

Figure 1. Faces that displayed the action units associated with the confused state

The group of Grafsgaard [11] conducted a study that aimed to examine how the context of dialogue and learning task were associated with student display of AU4. In Grafsgaard's study, participants solved an introductory computer programming problem and carried on computer-mediated textual dialogue with a human tutor. Facial recordings of students were collected using built-in webcams. The results indicated that students were significantly less likely to display AU4 (less likely to be confused) immediately following tutor questions, lukewarm feedback, and extra-domain dialogue acts, as well as during incomplete, on-track task actions.

The state of confusion was also observed among students who used an intelligent tutoring system designed to help learn arithmetic and algebra, *Aplusix II: Algebra Learning Assistant* [1]. In this study, actions, utterances, facial expressions, and body language, were used to code affect [28]. In *Aplusix*, topics are grouped into six categories (numerical calculation, expansion and simplification, factorization, solving equations, solving inequalities, and solving systems), with four to nine levels of difficulty each. The program presents the student with an arithmetic or algebraic problem from a problem set chosen by the student. Students then solve the problem one step at a time [2]. In this study, the confused student was described as *"someone who attempts a smaller number of problems and who works on a bigger number of easy problems compared to other students"* [1]. Moreover, students who got the lowest number of correct answers, attempted the less difficult problems, and took the longest time in solving algebra problems, experienced confusion [28]. Behaviourally, the confused student was observed as one who was scratching his/her head, repeatedly looking at the same interface elements; consulting with a fellow student or the teacher; looking at another student's work to determine what to do next; giving statements like, *"I'm confused!"* or *"Why didn't it work?"* [2].

The focus of this study is the confusion in the experience of the novice programmer. In 2011, Lee [17] created a model that could detect confusion among novice programmers. Through the model,

we learn that confusion usually occurs when the student is faced with an obstacle, normally an error, and that the confused novice programmer normally had compilation with errors more than half of the time. In a confused state, the student might take time before making the next compilation, and the consequent compilation would result in the same error.

In other words, the confusion we are interested in this study is one that drives a student to a seeming impasse when confronted with unexpected compiler messages, hence, displaying errors in his program, and resulting in distinct facial expressions and gestures normally attached to a confused person. We combine qualities and observations given above for the operational definition of confusion. Through the videos, we look for facial expressions that display facial action units (AU's) that are associated with confusion (AU4 - brow lowerer, AU7 – lid tightener, and lips or mouth that do not show AU12 – lip corner puller.) Then we check the corresponding programming screen interface – inactivity and uncertainty indicated by untouched programs, or interface other than the programs like Web sites for notes and other help sites, suggest confusion. Other activities like talking to peers, or looking at other note sources, would also indicate confusion.

3. METHODOLOGY

3.1 Participants
Twenty students enrolled in an introductory programming course (CS21A) in the first semester of SY 2008-2009 were invited to participate in an experiment designed to capture facial expressions and gestures while working on a Java computer programming exercise in the lab.

3.2 Sources of Data
There were three sources of data: video capture of the students while working on the program; video capture of the program; and compilation logs. The computer interface that the students used was video-taped, and the interaction of the students with the compiler was captured in a network database.

3.3 Data Cleaning
The following reasons resulted in reducing the data set from 20 to 12:

1. Each student must have two videos. However, four students had only one video, resulting in 16 videos.

2. The video capture for one student was not reliable because proper display of the AU's and gestures were not captured by the camera. The student either would cover her face or evade the camera. For this student, only seven out of forty-eight time slices could be coded. This reduced videos to 15.

3. Table 1 and Table 2 show the data for the 15 students initially considered in the study. Table 1 shows the data for each student from the compilation logs. Table 2 shows the data from the videos. We can see that we have 15 students initially. However, due to the following incidents, the final data set was reduced to 12. The student in video 104 had 19 time slices only. The student in video 107 did not have compilation logs. The student in video 203 had incomplete data, that is, he only had 13 time slices whereas his compilation

log showed he completed his task. Thus, we had 12 videos left to work with.

Table 1. Compilation log data of each student

N	Student in Video	No. of Compilations	No. of Errors	Average time between compiles in seconds	Confusion Rate from BlueJ	EQ
	From Compile Logs					
1	101	34	31	86.000	1	0.476
2	103	71	55	35.549	1	0.464
3	104	12	10	216.917	1	0.568
4	106	23	15	71.391	1	0.354
5	107					
6	108	20	9	109.550	0.5	0.183
7	109	15	5	86.867	1	0.111
8	110	29	10	56.448	1	0.386
9	201	59	42	37.458	0.57	0.364
10	202	28	16	40.250	0.67	0.415
11	203	48	45	50.750	1	0.501
12	301	49	11	57.204	0.33	0.063
13	302	83	78	42.349	1	0.732
14	303	42	38	49.976	1	0.604
15	308	15	15	58.333	1	0.484

Table 2. Video data of each student

N	Student in Video	No. Times Confused	No. Times Not Confused	% Confused	Total No. Time Slices
	From Videos				
1	101	46	5	90.2	51
2	103	34	21	61.82	55
3	104	15	4	78.95	19
4	106	34	8	80.95	42
5	107	36	5	87.8	41
6	108	35	15	70	50
7	109	33	10	76.74	43
8	110	31	10	75.61	41
9	201	24	29	45.28	53
10	202	23	18	56.1	41
11	203	8	5	61.54	13
12	301	49	7	87.5	56
13	302	60	5	92.31	65
14	303	38	7	84.44	45
15	308	58	4	93.55	62

3.4 Synchronization and annotation of the videos

The two videos were synchronized in ELAN and two coders used a manual in coding each 10-second of every minute, as "confused" or "not confused." See 3.5 for the criteria used on labelling video clips.

3.5 Data Coding

A manual was created as guide in coding each 10-second time slice. The criteria are enumerated below.

1. The presence of any of the following Facial Action Coding System units that have been identified in other studies to be indicative of confusion: lowered brows, tightened eyelids, and the lack of a lip corner puller.
2. Scratching of head
3. Looking at the same computer interface repeatedly
4. Consulting with fellow student or teacher
5. Staring at the screen without doing anything
6. Writing and then deleting characters/words on the screen
7. Mouse goes randomly on the screen
8. Consulting notes for the 10-second observation period
9. No change in the program for the 10-second observation period

The criteria given above were taken from related studies. Item 1 was largely based on studies that used Ekman's Facial Action Coding System [5, 11, 23]. Items 2, 3, and 4 were among the descriptors for a confused student learning Algebra by using *Aplusix* [1, 2, 28]. Items 5, 6, 7, 8, and 9 were behavioural or operational definitions of the definition of confusion from the studies of the group of D'Mello and Graesser, which was "noticeable lack of understanding and being unsure about how to proceed" [5, 6, 7, 8]. In judging the affect, the coders agreed that the presence of any of the enumerated criteria rendered the student "confused."

4. RESULTS

4.1 Correlation Analysis

We wanted to see whether there is a relationship between the rate of confusion from the video observations and each of the following data from the compilation logs: number of compilations, number of errors, total time taken by the student in working on the task, average time between compiles, confusion rate and error quotient or EQ. Correlation values for each of these compilation log data and rate of confusion or % confusion from video observations were computed. Table 3 shows these values as well as the significance p.

Table 3. Values for Correlation r and Significance p

From Video	From Compilation logs	r	p
% confusion	Number of Compilations	-.074	.819
% confusion	Number of Errors	.042	.897
% confusion	Total Time	.247	.439
% confusion	Average Time between Compiles	.281	.376
% confusion	Confusion Rate	.395	.204
% confusion	EQ	.214	.504

Confusion Rate and EQ values were lifted directly from the logs. The number of compilations, and number of errors, can be derived through inspection of the log table, while total time and average time between compiles were computed. From the video data, % confusion was computed as the result of dividing the number of time slices that were given the label "confused" by the total number of slices, which is the sum of time slices that were judged "confused" and "not confused." The following paragraphs present the individual correlation analysis for each feature of the compilation log and % confusion.

4.1.1 Number of Compilations and % confusion
The number of compilations was found to be not significantly related to % confusion at the .05 level, r=-.074, p=.819.

The significance value tells us that the probability of getting a correlation coefficient this big (actually a very small value of .074) in a sample of 12 if the null hypothesis were true (there was no relationship between these variables) is high (.819 or roughly 82%). This value is much greater than .05, the normal criterion for significance. Hence, the relationship between the number of compilations and % confusion cannot be established statistically. The negative sign for the Pearson correlation could have pointed to the direction of the relationship, that is, as one variable increases, the other one would decrease. On the other hand, even if the significance value were greater than .05, a Pearson coefficient of .074 would be interpreted as negligible [24], that is, if number of compilations increased, there would be almost no change in the value of % confusion [10].

4.1.2 Number of Errors and % confusion
The number of errors was found to be not significantly related to % confusion, r=.042, p=.897. If the significance value p were less than .05, we would say that the relationship between number of errors and % confusion is negligible (r=.042), that is, a change in the number of errors would have no effect on the % confusion.

4.1.3 Total time and % confusion
Total time was found to be not significantly related to % confusion, r=.247, p=.439. If the relationship had been significant, a correlation coefficient of .247 renders a small effect, or weak positive relationship [10].

4.1.4 Average Time and % confusion
The average time was found to be not significantly related to % confusion, r=.281, p=.376. The r value of .281 also gives a positive weak relationship, had the significance value been less than the criterion standard for significance (.05).

4.1.5 EQ and % confusion
Error Quotient (EQ) was found to be not significantly related to % confusion, r=.214, p=.504. The r value of .281 also gives a positive weak relationship, had the significance value been less than the criterion standard for significance (.05).

4.1.6 Confusion rate and % confusion
Confusion Rate was found to be not significantly related to % confusion, r=.214, p=.504. Looking at r alone (r=.395), the Confusion rate could have had a moderately strong relationship with % confusion. We can say that the significance value tells us that the probability of getting a correlation coefficient of .395 in a sample of 12 if the null hypothesis were true (there was no

relationship between these variables) is still high at .204. In other words there is a 20% chance that we shall get the Pearson coefficient of .395 even if here was no relationship between confusion rate and % confusion.

4.2 Discussion
Looking at the correlation values alone, the relationships between % confusion and number of errors (r=.042), and % confusion and number of compilations (r=-0.074), are negligible. This means that knowledge of number of errors, or of number of compilations, does not give us information about what the value of % confusion is likely to be, and vice versa. On the other hand, the relationship between Total Time and % Confusion (r=.247), average time between compiles and % confusion (r=.281), and % confusion and EQ (r=.214) is weak [24]. % confusion and confusion rate have a moderate positive relationship [24], meaning a high confusion rate may correspond to a high value for % confusion, and vice versa.

Note, however, that all the p values are greater than 0.05, meaning the results did not achieve statistical significance.

With the small sample used, it was difficult to get a powerful relationship between the variables we wanted to study. However, we still considered the relationships because this might change and become significant with a bigger sample. To explain how this can happen, refer to a Pearson's correlation table [26] that shows level of significance for one-tailed and two-tailed researches with alpha level of .05. The column with heading "Level of significance for two-tailed test" gives the critical values for Pearson's r that are needed to be surpassed to achieve significance. The study is two-tailed because we cannot predict the direction of the relationship, that is, for example, we do not know if increasing the number of compilations will result in increase or decrease in confusion [10]. Using the .05 level and travelling down the column we can see that as the sample size gets larger, the size of the correlation that is needed to achieve significance gets smaller. For a sample size 12 (df=10), the critical value is .576 (or -.576). The values of Pearson correlation coefficient we got in this study were lower than .576, hence these values are not significant at the .05 level. The biggest r we got, which was .395 for Confusion rate and % confusion, would have been significant if our sample size were at least 26.

Another reason we can look into as to the reason why the relationships were not significant is the premise of using Facial Action Coding System or FACS. Although FACS was created to prove that "basic emotions" such as anger, happiness, surprise, fear, disgust, and sadness are universal, it might just well be that the display of emotions are cultural. Matsumoto stated that there are display rules observed in different cultures [22]. Display rules that are learned early in life affect individuals in modifying or managing their emotions. For example, Japanese individuals may display more emotions in the presence of higher-status individuals. In our study, students were well-aware that they were being observed, and that "higher-status" individuals (teacher, for example) were present in the room. Perhaps we looked for facial expressions that were not true to our culture, or were actually subjected to display rules exercised in our culture. These are areas that can be explored further.

In this research, individual features of a student's compilation logs were correlated with what was perceived or defined to be a confused state, based on facial expressions and gestures. The operational definition of confusion was based on findings from

related studies, and was amplified by the author's interpretation of the textual definition of "confusion" from related studies, specifically, "noticeable lack of understanding and being unsure about how to proceed."

In Lee's study [17], where a model was made to determine confusion using compilation logs, all of 6 features from the compilation logs were used to build the model. The six features were: Average time between compilations, maximum time between compilations, average time between compilations with errors, maximum time between compilations with errors, number of compilations with errors, number of pairs of consecutive compilations with the same errors). It may be that the individual feature of the compilation logs was not enough to "create" a significant relationship with confusion observed from the videos. On how this will be done, what combinations of compilation logs features may be used, can be an area for future study.

Although the statistics do not give credence to the results, it is worth noting, for the purposes of further research, what we thought we found in this study. For the novice programmer, an impasse was exemplified by staring at the screen, with random mouse movements and scrolling up and down of the screen. Encountering an impasse results in cognitive disequilibrium and confusion has been described to accompany cognitive disequilibrium. Facial expressions observed during these moments showed hand and mouth expressions aside from what has been published in literature like occurrence of lowered brows and tightened eyelids. Hand positions like hands on the face (a hand or both hands), hand on the chin, hand on mouth, and mouth expressions like pursed, sucked, or opened/parted mouth were among the expressions observed during moments when students were judged confused.

The results may be improved by modifying the method of coding. First, instead of having only one set of data, we can include affect judgment by an expert, or someone who has done affect judgment through video before. Multiple judges are justified because "there is no clear gold standard to declare what the learner's states truly are" [9]. Moreover, according to D'Mello and Graesser [9] emotions do not last for more than 4 seconds, occurring from between .5 to 4 seconds. The affect judgments could have been done with smaller time slices, like every 3 seconds, or 4 seconds. We have actually encountered shifts in affect within the 10-second observation period. However, we relied on the facial expressions, and activity of the student, occurring for a longer time. If the student spent more time modifying or making significant modification in the program, the slice would be labelled "Not Confused"; if the video showed the student smiling, and busy looking at notes and the program he was working on, this would be labelled "Not Confused." However, if the student stopped, stared at the screen without making changes in the program, or showed hesitation by writing and deleting the same characters, this would be labelled "Confused."

5. CONCLUSION
In this study, we wanted to answer the following questions:
1. How can we quantify novice programmer confusion?
2. What is the relationship between confusion and compilation behaviour?

Quantification of programmer confusion based on facial expression, gestures and program screen output was made as the per cent times the student was confused, that is, the number of time slices judged as "confused" over all time slices. Confusion was operationally defined as a time when the novice programmer encounters an impasse exemplified by staring at the screen, with random mouse movements and scrolling up and down of the screen. Facial expressions observed during these moments showed hand and mouth expressions aside from what has been described in literature like occurrence of lowered brows and tightened eyelids.

In this study, the relationship between video rate of confusion and number of errors, video rate of confusion and number of compilations, video confusion rate and total time, video confusion rate and average time between compilations, video confusion rate and EQ, and video confusion rate and compiler confusion rate, did not achieve statistical significance.

Improvement on the coding method is recommended. First, instead of having only one set of data, we can have another set of data from other coders. Affect judgment by an expert, or someone who has done affect judgment before, may be included. Also, affect judgments may be done with smaller time slices, like every 3 seconds, or 4 seconds, for the complete duration of the video.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Anglo, E. A., and Rodrigo, M.M.T. 2010. Can Affect Be Detected from Intelligent Tutoring System data? – A Preliminary Study. ITS (Intelligent Tutoring Systems) 2010, Part II, LNCS 6095, V. Aleven, J. Kay, and J. Mostow, (Eds.).

[2] Baker, R.S.J.D., D'Mello, S. K., Rodrigo, M.M.T. and Graesser, A.C. 2010. Better to Be Frustrated than Bored: The Incidence, Persistence, and Impact of Learners' Cognitive-Affective States during Interactions with Three Different Computer-Based Learning Environments. International Journal of Human-Computer Studies archive Volume 68 Issue 4, April, 2010 Pages 223-241 Academic Press, Inc. Duluth, MN, USA.

[3] Brugman, H., and Russel, A. 2004. Annotating Multi-media / Multi-modal Resources with ELAN. Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004). Sannella, M. J. 1994. Constraint Satisfaction and Debugging for Interactive User Interfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.

[4] Craig, S. D., Graesser, A. C., Sullins, J. and Gholson, B. 2004. Affect and Learning: An Exploratory Look into the Role of Affect in Learning with Auto Tutor. Journal of Educational Media, Vol. 29, No. 3, October 2004.

[5] Craig, S., D'Mello, S, Witherspoon, A., and Graesser, A. 2007. Emote-Aloud during Learning with Auto Tutor: Applying the Facial Action Coding System to Cognitive-Affective States during Learning. Cognition and Emotion, Volume 22, Number 5, December 2007 , pp. 777-788(12).

[6] D'Mello, S., et al. 2012. Confusion Can Be Beneficial To Learning. Learning and Instruction, doi:10.1016/j.learninstruc.2012.05.003.

[7] D'Mello, S. and Graesser, A. 2011. The Half-Life of Cognitive-Affective States during Complex Learning. cognition & emotion, 25:7, 1299-1308.

[8] D'Mello, S.K., Lehman, B., and Person, N. 2011. Monitoring Affect States during Effortful Problem Solving Activities. International Journal of Artificial Intelligence in Education, 20(4), 361-389.

[9] Ekman, P., Friesen, W., And Hager, J. 2002. Facial Action Coding System: The Manual. URL= http://face-and-motion.com/dataface/facs/manual/Purpose.html#913668.

[10] Field, Andy. 2009. Discovering Statistics Using SPSS (Third Edition). SAGE Publications Limited, London.

[11] Grafsgaard, J. F., Boyer, K. E., Phillips, R., and Lester, J. C. 2011. Modeling Confusion: Facial Expression, Task, and Discourse in Task-Oriented Tutorial Dialogue. The 15th International Conference on Artificial Intelligence in Education (AIED 2011).

[12] Hughes, M.C., Jadud, M.C., and Rodrigo, M.M.T. 2010. String Formatting Considered Harmful For Novice Programmers. Computer Science Education, ISSN 0899-3408, Volume 20, Issue 3, September 2010.

[13] Jadud, M.C. 2005. A First Look at Novice Compilation Behaviour Using BlueJ. Computer Science Education, Vol. 15, No. 1, March 2005, pp. 25 – 40.

[14] Jadud, M.C. 2006. Methods and Tools for Exploring Novice Compilation Behaviour. International Computing Education Research (ICER'06), September 9–10, 2006, Canterbury, United Kingdom.

[15] Jenkins, T. 2001. The Motivation of Students of Programming. Graduate thesis University of Kent. URL = http://www.cs.kent.ac.uk/people/staff/saf/dc/portfolios/tony/doc/other/thesis.

[16] Kinnunen, P. and Malmi, L. 2006. Why Students Drop Out CS1 Course. Proceedings of the second international workshop on Computing Education Research ICER '06 (2006).

[17] Lee, D. M. C. 2011. Detecting Confusion Among Novice Programmers Using BlueJ Compile Logs. Graduate thesis, Ateneo De Manila University.

[18] Lehman, B. A., Matthews, M., D'Mello, S. K., and Person, N. 2008. What Are You Feeling? Investigating Student Affective States During Expert Human Tutoring Sessions. In Proceedings of the Ninth International Conference on Intelligent Tutoring Systems (pp. 50-59). Berlin, Heidelberg: Springer-Verlag.

[19] Lehman, B. D'Mello, S. and Graesser, A. 2012. Confusion and Complex Learning During Interactions with Computer Learning Environments. Internet and Higher Education, June 2012, 15(3):184-194. DOI: 10.1016/j.iheduc.2012.01.002.

[20] Lehman, B., D'Mello, S., and Person, N. 2008. All Alone With Your Emotions: An Analysis Of Student Emotions During Effortful Problem Solving Activities. URL = http://141.225.218.248/web-cslwebroot/emotion/files/lehman-affectwkshp-its08.pdf.

[21] Lewis, C. 2010. Attrition in Introductory Computer Science at the University of California, Berkeley. Technical Report No. UCB/EECS-2010-132 October 17, 2010. URL = http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-132.html.

[22] Matsumoto, David. 2006. Culture and Emotional Expression. Hong Kong Conference on Cultural Influences on Behavior. Hong Kong University of Science and Technology, Hong Kong, China.

[23] McDaniel, B., D'Mello, S., King, B., Chipman, P., Tapp, K., and Graesser, A. 2007. Facial Features for Affective State Detection in Learning Environments. CogSci2007 – 29th Annual Conference of the Cognitive Science Society (August 1 – 4, 2007).

[24] Pearson's r Correlation. URL= http://faculty.quinnipiac.edu/libarts/polsci/statistics.html.

[25] Pratt, M. K. 2012. 10 Hot Skills for 2013. Computer World. September 24, 2012, pp. 28 – 29.

[26] Price, Ian (Web Administrator). Web Stat. School of Psychology, University of New England. URL= http://www.une.edu.au/WebStat/unit_materials/c6_common_statistical_tests/pearsons_table.htm

[27] Rodrigo, M.M.T., Baker, R.S., Jadud, M.C., Amarra, A.C.M., Dy, T., Espejo-Lahoz, M.B.V., Lim, S.A.L., Pascua, S.A.M.S., Sugay, J.O., Tabanao, E.S. 2009. Affective and Behavioural Predictors of Novice Programmer Achievement. International Conference on Innovation and Technology in Computer Science Education, Haifa, Israel.

[28] Rodrigo, M.M.T., and Lagud, M.C.V. 2010. The Affective and Learning Profiles of Students Using an Intelligent Tutoring System for Algebra. URL = http://curry.ateneo.net/~didith/2010AffectiveAndLearningProfiles.pdf

[29] Teague, D. 2009. A People-First Approach to Programming. Proceedings of the 11th Australasian Computing Education Conference (ACE 2009).