The Affective and Learning Profiles of Students Using an Intelligent Tutoring System for Algebra

Maria Carminda V. LAGUD[1], Ma. Mercedes T. RODRIGO[2]

[1] *Education Department, Ateneo de Manila University*
[2] *Department of Information Systems and Computer Science, Ateneo de Manila University*
minminvl@gmail.com, mrodrigo@ateneo.edu

Abstract. We investigate whether high-performing students' experience of affect (boredom, confusion, delight, flow, frustration, neutrality and surprise) is different from that of low-performing students while using Aplusix, an ITS for Algebra. We found that students with the highest number of correct answers experienced flow the most while students with the lowest number of correct answers experienced confusion and boredom the most. Students who attempted the most difficult problems experienced flow the most while students who tried the lowest levels experienced more boredom and confusion. Students who took the longest time in solving the algebra problems experienced confusion the most while students who took the shortest time experienced confusion the least. Students who used the most number of steps to solve a problem experienced confusion and boredom the most. Students who used the least number of steps experienced more flow.

Keywords: affect, Aplusix, affective profile, learning profile

1. Introduction

Aplusix II: Algebra Learning Assistant is an intelligent tutoring system (ITS) for algebra. The system allows a student to solve an algebraic equation on a step-by-step basis. At each step, Aplusix provides the student with feedback indicating whether indeed a prior step and a current step were mathematically equivalent. Errors are immediately visible, prompting the student to make corrections early in the solution process. Prior research on Aplusix has shown that it has the capability to increase learning by 70% to 250% [2].

As with many other ITSs, Aplusix tends to track student performance, using this as basis for its subsequent interactions with the student. However, learning is not just a cognitive process. Learning also involves affect. Piaget (1989, in [8]) wrote that there is no cognitive mechanism without the affective element since affectivity motivates the intellectual activity. Thus, in recent years, more and more researchers and educators have been studying affect and its relationship with learning.

Affect pertains to a broad class of mental processes, including feelings, emotions, moods, and temperament (Dictionary of Psychology, Second Revised Edition., s. v. "affect"). Affect is related to motivation in that learners have feelings, emotions, and moods that they bring to bear on a task. "Students are more motivated when they feel optimistic about their goals and the chances of meeting them and when students are more excited after success, they are more willing to engage in the behavior again" [10]. Affect is also related to learning and cognition. The more emotional students feel about a piece of material, the more likely they are to remember it [10].

We aim to describe the relationship between the affective and learning profiles of students while interacting with Aplusix. We use the term "affective state" to refer to one emotion, feeling, or mood that a student displays during an observation, while term "affective profile" is the percentage of time a student displays each affective state during an observation period. For the purposes of this study, focus will be given to the affective states boredom, frustration, confusion, flow and delight, following the research of Craig et al. [3] and Rodrigo et al [13, 14, 15]. We use the term "learning profile" to refer to the number of correct items the student solved, the highest difficulty level he or she attempted, his or her average time to solve an item, and his or her average number of steps taken to solve an item.

2. Methods

2.1 Research Subjects and Setting

The participants for the study were first and second year high school students from four schools within Metro Manila and one school in Cavite. The students' age ranged from 12 to 15 with an average age of 13.5 and a modal age of 14. One hundred and forty students participated in the Aplusix study (83 female, 57 male). They were all computer-literate but none of them had previously used Aplusix. In groups of 10, participants were asked to use Aplusix for 42 minutes.

2.2 Aplusix

Aplusix covers six topics or categories of algebra: numerical calculation, expansion and simplification, factorization, solving equations, solving inequalities and solving systems of equations or inequalities. Each of these categories is again broken down into four to nine levels of difficulty. At the start of an Aplusix tutorial session, the student firsts select a problem set. The ITS presents a problem that the student must solve. Using an advanced editor of algebraic expressions (See Figure 1), the student then makes step-by-step calculations towards the solution. Two black parallel bars between two steps mean that the two steps are equivalent. Two red parallel bars with an X mean that the two steps are not equivalent. Aplusix also gives reports on the student's progress on the resolution of the problem. The report may include existence of errors or of an expression not yet in its simplest form. At any time, the student can end the exercise, ask for a hint or for the final answer or solution to be shown [9].

2.3 Quantitative field observations

We collected data regarding student affective states using the quantitative field observation methods discussed in Baker, et al [1]. The observations were carried out by a team of six observers, working in pairs. The observers were Masters students in Education or Computer Science, and all but one had prior teaching experience. The observers trained for the task through a series of pre-observation discussions on the meaning of the affective categories. Observations were conducted according to a guide that gave examples of actions, utterances, facial expressions, or body language that would imply an affective state, and practiced the coding categories during an unrelated observation prior to this study.

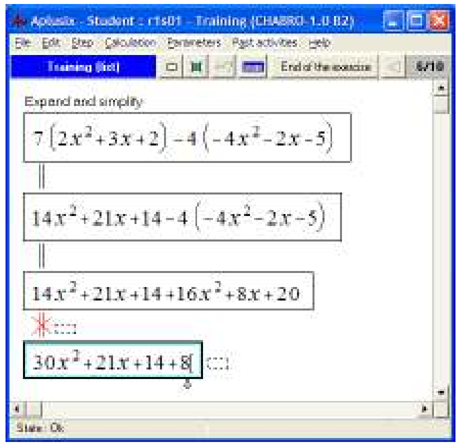

Figure 1. A screen shot of Aplusix

Each observation lasted twenty seconds. Observers stood diagonally behind or in front of the student being observed and avoided looking at the student directly, in order to make it less clear exactly when an observation was occurring. If two distinct affective states were seen during an observation, only the first was coded, and any behavior by a student other than the student currently being observed was not coded. Since each observation lasted twenty seconds, each student was observed once per 180 seconds. The affective categories coded used were those from Rodrigo, et al.'s [15] study and were boredom, confusion, delight, surprise, frustration, flow [5] and the neutral state. "Flow" refers to full immersion in an activity; the participant is focused on a task to the point that he or she is unaware of the passage of time [5]. The utterances, postures, and facial expressions that were used to identify each state were discussed in [15].

Some of these affective categories may not be mutually exclusive (such as frustration and confusion), though others clearly are (delight and frustration). For tractability, however, the observers only coded one affective state per observation. Thirteen pairs of observations were collected per student. Inter-rater reliability was acceptably high: Cohen's κ=0.63.

2.4 Aplusix log files

As the student uses Aplusix, the software logs all user interactions. Each exercise is recorded in one log.

Figure 2 shows how a raw log file looks like. The raw log files were processed, combined and summarized into a more comprehensible master table of log files, a part of which is shown in Table 1.

```
;FeuilleDeCalculs;2;actions;□{[Type FeuilleDeCalculs] [Activite Entrainement] [expression <<{[{[5*4*2]_etape_ □["Calculate"]_etape_
["40"]_etape_ ["TpbCalculer"]}]_donne_ [{[40]_etape_ [""]_etape_ □["Solved"]_etape_ [""]}]}>>] [score 1.000000000] [ExpressionGlobale
<<{[5*4*2]_etape_ □["Calculate"]_etape_ [""]_etape_ ["TpbCalculer"]}>>] [CreditsIndications -1]}□□
%;FICHIER-ACTIONS;#VersionProtocoles=1;#Appli=APLUSIX II1.01b□□%;ELEVE□#identifiant="r1s01"□#passe="0F/D2F-C0C"□#nom=""□
#prenom=""□#NumeroAnonyme=1□#Role=Eleve□#SousRole=()□#Email=""□#Langue=()□#premierlancement="false"□#RepSaveAs="C:"□
#laclasse="AHS Run 1"□#CheminRepertoireActivite="C:\PROGRAM FILES\APLUSIX II\APLUSIXSYSTEM\..\..\..\WINDOWS\All Users\Applicatior
Data\Aplusix\Classes\AHS Run 1\r1s01\D-2007-03-01\A-15-46-17-Entrainement"□#Activite=Entrainement□#DureeDuTest=1800□
#SorteListeExercice=ListeExercices□#ListeExercices=((<<5*4*2>> PatronCalculCE22) (<<-5-9+8>> PatronCalculCE12) (<<3(2-3)>>
PatronCalculCE31) □(<<4*2*5>> PatronCalculCE22) (<<6+7-5>> PatronCalculCE12) (<<3(2-3)>> PatronCalculCE31) □(<<5(3-5)+2(2-4)>>
PatronCalculCE32) (<<15-(-6-3)>> PatronCalculCE12) (<<-14*9>> □PatronCalculCE21) (<<15+7>> PatronCalculCE11))□#NumeroExercice=
#TempsTestRestant=0□#FichierListeExercices=""□#InfoListeEngendree=("CHABRO-1.0" "A1")□#ExerciceHorsListe=()□#FichierPointAlg=""□
#FeuilleCalculTerminee=vrai□#DocumentModifie=()□#verification=permanente□#LeScoreMax=20□□%;PARAMETRES□
#VerificationCalculs=ChoixEleve□#PorteeVerification=tout□#RaisonnementCorrect=non□#SansMalFormee=oui□#SansErronee=oui□
#CommandeCalculsNumeriques=decimal□#CommandeDevelopperReduire=non□#CommandeFactoriser=rien□#CommandeResoudreEquation=ri
#OrdreExerciceAleatoire=oui□#IncitationCommentaire=non□#PresentationSolution=oui□#RetourEnArriere=oui□#DureeTest=@30□
#petiteFleche=ChoixEleve□#EtapeInterneModifiable=oui□#AvecSelection=oui□#FrappeParentheses=texte□#FrappeFraction=infixee□
#CreationEleve=oui□#ScoreMaxi=@20□#PositionIncitationCommentaire=Partout□#VerificationResolu=vrai□□%;CHAMPS□
No;duree;action;erreur;etape;expression;etat;curseur;selection;equivalence;resolu;□□
%;ACTIONS;#Date=3/1/07#Heure=15:46:17;#TypeProbleme=TpbCalculer□0;0.0;structure;();0;();();();();();();□1;0.0;enonce;();0;5*4*2
devant);rien;;N1;□2;6.7;placerCurseur;();0;5*4*2;();(2 0 derriere);rien;;N1;□3;6.8;commentaireetape : 40;();0;5*4*2;();(2 0
derriere);rien;;N1;□4;11.8;dupliquer;();1;5*4*2;();(2 0 derriere);rien;V1;N1;□5;1.6;BackSpace;();1;5*4*?;();(2 dedans);rien;V-;N-;□
6;0.2;BackSpace;();1;5*4;();(1 0 derriere);rien;V0;N0;□7;0.2;BackSpace;();1;5*?;();(1 dedans);rien;V-;N-;□8;0.2;BackSpace;();1;5;();
```


Figure 2. Sample Raw Log File

Column No.	1	2	3	4	5	6	7	8	9	10
Heading	School	Run	Student No.	Set No.	Problem No. within the set	Absolute Problem No.	Date	Time Started	Level	Step No.
Rows of Data	Alphonsus	1	07	1	01	1	1/22/07	00:33:29	B1	0
	Alphonsus	1	07	1	01	1	1/22/07	00:33:29	B1	0
	Alphonsus	1	07	1	01	1	1/22/07	00:33:29	B1	0

continuation

Column No.	11	12	13	14	15	16	17	18	19	20
Heading	Duration	Action	Error	Etape	Expression	Etat	Cursor	Selection	Equivalence	Resolution
Rows of Data	0	structure	()	0	()	()	()	()	()	()
	0	enonce	()	0	-8x(-8x-4)	()	(devant)	rien		N1
	0	termine	()	0	-8x(-8x-4)	()	(devant)	rien	()	N1

Table 1. Sample Data from the Master Table of the Log Files

The columns in the table are as follows:

1. School – the name of the participating school
2. Run – the student's run or batch number. Three to four batches of 10 students each were observed per school. The run number ranges from 1 to 4.
3. Student No. – the identification number of the student within the run taking the exercise. The student number ranges from 1 to 10.
4. Set No. – the set number of the current exercise. A set is composed of a group of items under a specific exercise category
5. Problem No. Within Set – the item number within the set number chosen
6. Absolute Problem No. – the item no. relative to all the problems answered by the student
7. Date – the date when the exercise was done.
8. Time Started – the time when the student started with the specific problem (identified by column 5 or 6)
9. Level – the degree of difficulty of the topic.
10. Step No. – the number of the current step
11. Duration – the number of seconds describing how long each step was done

12. Action – the action performed by the student. Terms used are expressed in French (Fr.).
13. Error – the error committed by the student while solving the problem
14. Etape (Fr.) – the stage or phase of the solution
15. Expression – the state of the mathematical expression
16. Etat (Fr.) – the current state or condition of the solution
17. Cursor – location of the cursor. Examples for Cursor values are: devant (Fr.)- in front of/ outside, dedans (Fr.) – inside
18. Selection – selected values in the solution.
19. Equivalence – indicates whether the equation is correct or not.
20. Resolution – indicates if the problem has been solved or not.

Only column numbers 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 20 were used for this research to help identify the learning profile of the students.

2.4 Affective Profile.
The affective profile of each student was derived based on the quantitative field observations.

To generate the affective profile from the data collection instrument, points were given to affective states observed. An affective state noted by both observers will be given 1 point for the particular time slice. If the two observers did not agree, 0.5 point will be given for each of the observed affective state. The number of times when each of the affective states was observed will be divided by 13 (for the 13 time slices) to get the percentage of time an affective state has been observed from a student. Table 2 shows the point system used. From the table, Student ABC can now be described as being in the state of flow 84.62% of the time, confused 11.54% of the time and showed neutrality 3.85% of the time.

Time	Rater 1	Rater 2	BOR	CON	DEL	FLO	FRU	SUR	NEU	Total
1	NEU	FLO				0.5			0.5	
2	FLO	FLO				1				
3	FLO	FLO				1				
4	FLO	FLO				1				
5	FLO	FLO				1				
6	FLO	CON		0.5		0.5				
7	FLO	FLO				1				
8	CON	CON		1						
9	FLO	FLO				1				
10	FLO	FLO				1				
11	FLO	FLO				1				
12	FLO	FLO				1				
13	FLO	FLO				1				
Total			0 or 0%	1.5 or 11.54%	0 or 0%	11 or 84.62%	0 or 0%	0 or 0%	0.5 or 3.85%	13 or 100%

Table 2. Affective Profile of a Student ABC

2.5 Learning Profile.

Using the Aplusix logs, we determined each student's learning profile, defined as the number of problems correctly solved, the highest difficulty level attempted, the average time to solve a problem, and the average number of steps used to solve the problem.

The means and standard deviations of each of the four variables were computed. For each of the four variables, we grouped the students into three: those within one standard deviation (average group), those above one standard deviation (above average group) and those below one standard deviation (below average group). We found, though, that the above average and below average groups were very small, e.g. less than 10 people, as opposed to over 95 people in the average group. We therefore decided to group the students terciles or by dividing the sample into three groups (average, above average, below average) centered on the median.

2.6 Comparison

We then computed for the affective profile of each tercile by taking the average incidence of each affective state of each student within the group. The affective profiles of the terciles were then compared with one another using a One-Way Analysis of Variance (ANOVA) with Statistical Package for Social Sciences (SPSS). Table 3 shows the tercile where the same Student ABC used in Table 2 belongs.

Student	BOR	CON	DEL	FLO	FRU	SUR	NEU	Total
AAA	0.00%	23.08%	7.69%	69.23%	0.00%	0.00%	0.00%	100.00%
AAB	0.00%	7.69%	0.00%	92.31%	0.00%	0.00%	0.00%	100.00%
ABB	0.00%	3.85%	0.00%	96.15%	0.00%	0.00%	0.00%	100.00%
ABC	0.00%	11.54%	0.00%	84.62%	0.00%	0.00%	3.85%	100.01%
BBC	7.69%	3.85%	3.85%	76.92%	3.85%	0.00%	3.85%	100.01%
:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:
YYZ	0.00%	7.69%	0.00%	61.54%	30.77%	0.00%	0.00%	100.00%
XYZ	0.00%	3.85%	11.54%	84.62%	0.00%	0.00%	0.00%	100.01%
47 students	4.75%	16.53%	5.89%	68.49%	2.70%	0.33%	0.00%	99.67%

Table 3. Sample Tercile Group

3. Results and Discussion

The following ANOVA results were obtained. Scheffe posthoc tests were also done to identify which particular groups showed significant differences.

In terms of correct items solved, the group that scored the highest (above average) experienced flow the most (F=3.948; p=0.022) with a mean difference of ± .106014 between the above and below average groups from the Scheffe posthoc tests. The group that scored the lowest experienced (below average) the most boredom (F=3.995; p=0.021) with a mean difference of ± .040057 between the above and below average groups. The group that scored the lowest experienced the most confusion (F=5.163; p=0.007) with a mean difference of ± .070853 between the above and below average groups..

In terms of highest difficulty level attempted, the group that tried the most difficult levels (above average) experienced flow the most (F=5.994; p=0.003) with a mean difference of ± .125430 between the above and below average groups. The group that tried the lowest levels experienced significantly more boredom (F=5.495; p=0.005) with a mean difference of ± .045826 between the below average and average groups and a mean difference of ± .042088 between the above and below average groups. Same with boredom, the group that tried the lowest levels experienced significantly more confusion (F=6.006; p=0.003) with a mean difference of ± .073079 between the above and below average groups.

The group that took the longest time in solving (below average) experienced confusion the most while the group that took the shortest time in solving (above average) experienced confusion the least (F=4.726; p=0.010) with a mean difference of ± .064378 between the above and below average groups.

Finally, the group that used the most number of steps (below average) experienced confusion (F=4.082; p=0.019) with a mean difference of ± .057691 between the above and below average groups. The below average group also experienced boredom the most (F=3.617; p=0.029) with a mean difference of ± .040382 between the average and below average groups. The average group and the group that used the least number of steps experienced flow more than the group that used the most number of steps (F=3.476; p=0.034).

All learning groups in all categories experienced flow to a great extent. The significant differences lie in the degree to which the groups experienced flow. The groups that scored the highest based on the number of correct items solved, those who tried the most difficult levels and those who took fewer number of steps in solving an item experienced significantly more flow than those in other groups. These findings support findings from other studies that indicate that flow is experienced more by people who are more motivated, those who are willing to go further, those willing reach for higher levels of challenge and are achievers or experts [4, 5, 12]. There are also significant differences in the degree to which the learning groups experienced confusion. The group that scored the lowest, the group that answered items in the lowest levels, the group that took the most number of steps in solving an item and the group that used the most time in answering an item experienced significantly more confusion than other groups. These are aligned findings from definitions of confusion: a feeling of perceptual disorientation and lack of clear thinking [6] or a feeling of not knowing, when information is not present in memory [7]. On the other hand, confusion can also be equated with a constructive form of cognitive dissonance, which is positively related to optimum learning gains [3]. This may account for the fact that groups with higher levels achievement experienced confusion as well.

In terms of boredom, the group that scored the lowest, the group that used the most number of steps and those who stayed in the lowest difficulty level experienced significantly more boredom the other groups. The students may have used repetitive steps that contributed to the high quantity of steps used. For one, based from the logs, the student who used the highest average of steps (i.e. 403 steps) only tried answering three items. This student tended to type numbers and erase them repeatedly—as if he was thrashing. According to English & English [6] boredom is felt when doing uninteresting activities. Perkins and Hill [11] also discussed the association of boredom with subjective monotony.

4. Conclusion

The results of this study are consistent with intuition: Good students tend to experience more flow, less boredom and less confusion than students who are struggling. What makes this study interesting, though, is that it attempts to quantify levels of achievement and their associated affective states. In the design of affective interventions, findings such as these might give ITS designers clues as to how students who are performing poorly are feeling. Designers might therefore arrive at appropriate intervention strategies that address not just the students' cognitive problems but their affect-related issues as well.

Acknowledgements. We thank Dr. Ryan Baker for his continuing support and guidance. We thank Jean-Francois Nicaud of the Laboratoire d'Informatique de Grenoble for the use of Aplusix. We thank Sheryl Ann Lim, Alexis Macapanpan, Sheila Pascua, Jerry Santillano, Jessica Sugay, Sinath Tep, and Norma Jean Viehland, and Dr. Ma. Celeste T. Gonzalez for their assistance in organizing and conducting the studies reported here. We also thank the Ateneo de Manila High School, Kostka School of Quezon City, School of the Holy Spirit of Quezon City, St. Alphonsus Liguori Integrated School and St. Paul's College Pasig for their participation in the studies conducted. This research undertaking was made possible by the Philippines Department of Science and Technology Engineering Research and Development for Technology Consortium under the project "Multidimensional Analysis of User-Machine Interactions Towards the Development of Models of Affect".

References

1. Baker, R.S., Corbett, A.T., Koedinger, K.R., and Wagner, A.Z.: Off-Task Behavior in the Cognitive Tutor Classroom: When Students "Game the System". In: Proceedings of ACM CHI 2004: Computer-Human Interaction, pp. 383-390 (2004)
2. Bouhineau D., Nicaud J.F., Chaachoua H., Huguet T., Bittar M., Bronner A.: Two years of use of the Aplusix System. In: Proceeding of the 8th IFIP World Conference on Computer in Education. University of Stellenbosch, Cape Town South Africa, p. 23 (2005)
3. Craig, S. D., Graesser, A. C., Sullins, J., Gholson, B.: Affect and Learning: An Exploratory Look into the Role of Affect in Learning with AutoTutor. Journal of Educational Media, vol. 29, No. 3, October 2004. pp. 241-250. The University of Memphis (2004)
4. Csikszentmihalyi, M., Rathunde, K., Whalen, S.: Talented Teenagers: The Roots of Success and Failure. Cambridge University Press, New York (1993)
5. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper & Row, New York (1990)
6. English and English: A comprehensive dictionary of psychological and psychoanalytic terms. New York: Longmans, Green and Co. (1958)
7. Hess, U.: Now You See It, Now You Don't - The Confusing Case of Confusion as an Emotion: Commentary on Rozin and Cohen. Emotion, 3(1), pp. 76-80 (2003)
8. Jaques, P. A., Bocca, E., Vicari, R. M.: Considering Student's Emotions in Computational Educational Systems. In: Simposio Brasileiro de Informatica na Educacao, pp. 543-552. UFRJ: Rio de Janeiro (2003)

9. Nicaud J.F., Bouhineau D., Chaachoua H., Huguet T. and Bronner A.: A computer program for the learning of algebra: description and first experiment. In Proceedings of PEG2003. Saint Petersburg. (2003)
10. Ormrod, J.: Educational Psychology: Developing Learners. Sixth Edition. Pearson Merril Prentice Hall, New Jersey (2008)
11. Perkins, R. E. and Hill, A. B.: Cognitive and affective aspects of boredom. In: British Journal of Psychology, 76, pp. 221--234 (1985)
12. Rathunde, K., Csikszentmihalyi, M.: Middle School Students' Motivation and Quality of Experience: A Comparison of Montessori and Traditional School Environments. American Journal of Education 111(3), pp. 341-371 (2005)
13. Rodrigo, M. M. T., Anglo, E. A., Sugay, J. O., Baker, R. S. J. d.: Use of Unsupervised Clustering to Characterize Learner Behaviors and Affective States while Using an Intelligent Tutoring System . International Conference on Computers in Education 2008, In: Chan, T.-W., Biswas, G., Chen, F.-C., Chen, S., Chou, C., Jacobson, M., Kinshuk, Klett, F., Looi, C.-K., Mitrovic, T., Mizoguchi, R., Nakabayashi, K., Reimann, P., Suthers, D., Yang, S., & Yang, J.-C. (Eds.). International Conference on Computers in Education 2008, pp. 57-64. Asia Pacific Society for Computers in Education (2008)
14. Rodrigo, M. M. T., Baker, R. S. J. d., D'Mello, S., Gonzalez, M. C. T., Lagud, M. C. V., Lim, S. A. L., Macapanpan, A. F., Pascua, S. A. M. S., Santillano, J. Q., Sevilla, L. R. S., Sugay, J. O., Tep, S., Viehland, N. J. B.: Comparing Learners' Affect While Using an Intelligent Tutoring System and a Simulation Problem Solving Game. In: B. Woolf et al. (eds.): ITS 2008. LNCS 5091, pp. 40-49, Springer-Verlag, Berlin Heidelberg (2008)
15. Rodrigo, M. M. T., Baker, R. S. J. d., Lagud, M. C. V., Lim, S. A. L., Macapanpan, A. F., Pascua, S. A. M. S., Santillano, J. Q., Sevilla, L. R. S., Sugay, J. O., Tep, S., & Viehland, N. J. B.: Affect and Usage Choices in Simulation Problem-solving Environments. In: R. Luckin, K. R. Koedinger, J. Greer (Eds.), Artificial Intelligence in Education: Building Technology Rich Learning Contexts that Work. IOS Press, Amsterdam, The Netherlands pp. 145-152 (2007)