Woolf, B.P., Arroyo, I., Muldner, K., Burleson, W., Cooper, D., Dolan, R., Christopherson, R.M., The Effect of Motivational Learning Companions on Low-Achieving Students and Students with Learning Disabilities, International Conference on Intelligent Tutoring Systems, 2010, Pittsburgh.

The Effect of Motivational Learning Companions on Low Achieving Students and Students with Disabilities

Beverly P. Woolf,[1] Ivon Arroyo,[1] Kasia Muldner,[2] Winslow Burleson,[2]
David Cooper,[1] Robert Dolan,[3] Robert M. Christopherson [2]

[1] University of Massachussetts Amherst {ivon, bev, dcooper}@cs.umass.edu
[2] Arizona State University
{winslow.burleson, Katarzyna.Muldner, robert.christopherson@asu.edu}
[3] Pearson Education Bob.Dolan@Pearson.com

Abstract. We report the results of a randomized controlled evaluation of the effectiveness of pedagogical agents as providers of affective feedback. These digital learning companions were embedded in an intelligent tutoring system for mathematics, and were used by approximately one hundred students in two public high schools. Students in the control group did not receive the learning companions. Results indicate that low-achieving students—one third of whom have learning disabilities—had higher affective needs than their higher-achieving peers; they initially considered math problem-solving more frustrating, less exciting, and felt more anxious when solving math problems. However, after they interacted with affective pedagogical agents, low-achieving students improved their affective outcomes, e.g., reported reduced frustration and anxiety.

Keywords: Affective feedback, pedagogical agents, special needs populations.

1 Introduction

Effective teachers regularly address students' emotional states and social backgrounds [1]. If tutoring systems are to interact naturally and supportively with students, they need to provide an environment that recognizes affect and expresses socio-emotional competence to address affective challenges and fluctuations in individual affective states. In recent years, researchers have made significant improvements in modeling students' affect [2, 3, 4, 19]. While progress has been made, very little empirical research has been conducted on how digital learning environments should respond to individual students' affect and how differences among students impact this process; yet for exceptions, see [5, 6].

Within digital learning environments, animated pedagogical characters have the potential to support students by engaging them through social interaction. Up until now, the use of pedagogical agents has mainly focused on the cognitive rather than affective aspects of learning [7]. While some effort has been made to create affective agents [8, 9], evaluation of their impact in schools is still preliminary.

Here, we report on an evaluation of pedagogical agents with about 100 students in two rural, public high schools in the northeastern U.S. We focus on the impact of affective learning companions on low achieving students (including ones with disabilities) and begin with a description of how such students need affective support when learning math. We then describe the test bed tutoring system, the learning companions, and the experiments. We present the results and conclude with a discussion of implications for intelligent tutoring systems.

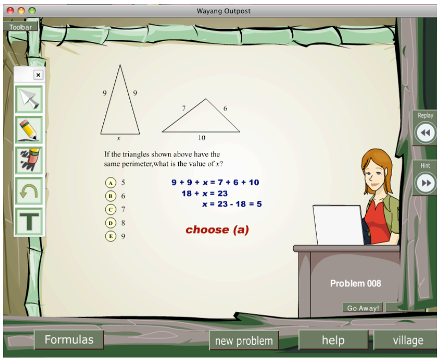

Fig. 1. The Wayang Tutor with Jane, the female affective learning companion.

2 Learning Disability and Low Achieving Students: Affective Needs

Classroom interventions (e.g., providing extra time on tasks, peer tutoring) that are effective for students with learning disabilities (LD) are difficult or impossible to sustain in classrooms without additional instructional support, something that schools are increasingly unable to provide due to budgetary constraints. Currently, students with learning disabilities who require extra resources comprise 13% percent of students in USA and 6.5% of the Massachusetts school population [10,11]. To the extent that these students are not being educated to their full potential, there is a large negative impact not only in the lives of these students but on society at large.

The under-achievement of students with LD in math does appear to have a biological basis, and there is evidence that many of these students have difficulties with working memory, executive control and procedural knowledge [12, 13]. As a result, many students with LD may persist in using counting strategies (e.g., finger counting) long after their typically achieving peers have switched to retrieving answers from memory [14], taking longer to solve math problems and performing poorly in math class and high-stake tests [15]. Students with LD develop more negative feelings towards math, choose less advanced math classes in high school and are later under-prepared for science and math careers. LD is a complex multi-factor problem and most educational institutions do not have the tools needed to provide cost-effective instruction tailored to each individual.

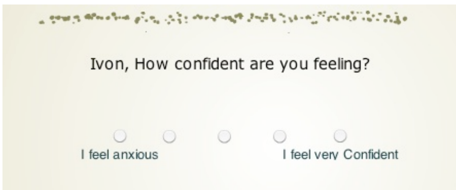

Fig. 2. Student's emotion self-reports within the tutor

Since low achieving students (both with and without disabilities) struggle with math, our conjecture was that *all* low achievers could require additional affective support. Thus, the first goal of the current study was to examine the affective needs of both low achieving and learning disability students in our data (15% of subjects). For the purpose of this paper we did not separately analyze differences between low-achieving and learning disability, because as a starting point we wanted to analyze what kind of support all low achievers require.

Table 1. Affective self-reports of high-achieving vs. low-achieving students prior to tutoring.

Affective Criterion	**Means, standard deviations and between-subjects test** Low-achieving: N=64; High-achieving: N=43
Self-concept of math ability (in comparison to other students, other subjects, 3 items)	Low-achieving: M=3.2 SD=1.1 High-achieving: M=4.1 SD=1.0 *****F(106,1)=18.2, p=.000**
How confident do you feel when solving math problems?	Low-achieving: M=3.1 SD=1.3 High-achieving: M=4.0 SD=1.3 *****F(105,1)=11.5, p=.001**
How frustrating is it to solve math problems?	Low-achieving: M=3.6 SD=1.2 High-achieving: M=3.0 SD=1.1 **** F(106,1)=7.6, p=.007**
How exciting is it to solve math problems?	Low-achieving: M=2.2 SD=1.2 High-achieving: M=2.7 SD=1.4 ***F(106,1)=3.64, p=0.05**

Data was collected and mean differences analyzed from a series of affective pretest questions given to students before tutoring, Table 1. The pretest covered general attitudes towards math and learning, such as likes/dislikes of math, how much was math valued as important, and how students felt when they solved math problems (anxiety, confidence, frustration, boredom, excitement). Low-achieving students were defined as those who scored lower than median grade on the math pretest. One third of these low-achieving students had been previously diagnosed as having a specific learning disability in math or reading and had an Individualized Education Plan (IEP), a document that identifies a student's academic, physical, social and emotional needs. Most students with IEPs (95%) are part of this low-achieving group. Table 1 shows that low-achieving students disliked math more, valued it less, had worse perception of their math ability, and reported feeling worse when solving math problems. We now present our test bed application, Wayang Outpost.

3. The Testbed Tutoring System: Wayang Outpost

Wayang Outpost ("Wayang") is an intelligent tutor that helps students prepare for standardized tests that assess general mathematic skills, see Figure 1 [17]. Problems are presented one at a time; each problem consists of the problem statement with four or five solution options directly below it. Students select an answer and the tutor provides immediate visual feedback by coloring the answer green or red, for correct or incorrect respectively. Prior to or after selecting an answer, a student may ask for a hint, which Wayang displays in progression from general suggestions to the correct answer. In addition to this domain-based help, Wayang includes a wide range of meta-cognitive and affective support, delivered by learning companions; agents designed to act like peers who care about a student's progress, and offer support and advice on how to improve student learning strategies. Wayang includes gendered and ethnically different companions allowing us to explore how the gender and the ethnicity of the companion influences outcomes (e.g., learning, attitudes) [18]. The learning companions' interventions are tailored to a given student's needs according to Wayang's affect and effort models embedded in the tutor. The *effort* model provides information on the degree of effort a student invests in generating a problem solution. A linear regression *affect* model is used to assess a student's emotional state; this model is derived from data obtained from a series of studies described in [18, 19].

4. Affective Support delivered by Wayang's Learning Companions

Learning companions deliver approximately 50 different messages emphasizing the malleability of intelligence and the importance of effort and perseverance (Table 2). The messages also include meta-cognitive help related to effective strategies for solving math problems and effective use of Wayang's tools. Ultimately, the interventions will be tailored according to Wayang's affective student model. However, we are currently still validating the models and algorithms for deciding which intervention to provide and when, and thus relied on the effort model only to assign messages for this experiment. This section describes these interventions including *attribution* and *strategy* training, as well as *effort affirmation*.

The affective support provided by Wayang in this experiment was to train students motivationally, by emphasizing the importance of effort and perseverance and the idea that intelligence is malleable instead of a fixed trait [20]. The characters provided this support by responding to the effort exerted by students rather than to the student's emotions. Characters were either unimpressed when effort was not exerted, or simply ignored that the student solved the problem. They also offered praise to students who exerted effort while problem-solving, even if their answers were wrong, highlighting that the goal is to lessen the importance of performance in favor of learning.

The characters were highly positive, in the sense that they displayed encouraging gestures (e.g., excitement and confidence). In a separate completed study, which is beyond the scope of this paper, characters behaviorally mimicked student self-reported emotions, which is a form of a non-verbal empathetic response (e.g., learning companions appeared excited in response to student excitement, see Figure 2, right). In this experiment reported here, the companions occasionally expressed non-verbal behaviors of positive valence only, the underlying goal being to make them appear life-like and engaged, and to impart some of their enthusiasm to the students. The next three types of interventions described are verbal messages tailored according to Wayang's modeling of students' effort.

Table 2. Companions provided several responses based on student effort

Type	Sample message
Attribution (General)	I found out that people have myths about math, thinking that only some people are good in math. Truth is we can all be good in math if we try.
Attribution (Effort)	Keep in mind that when we are struggling with a new skill we are learning and becoming smarter!
Attribution (No Effort)	We will learn new skills only if we are persistent. If we are very stuck, let's call the teacher, or ask for a hint!
Attribution (Incorrect)	When we realize we don't know why the answer was wrong, it helps us understand better what we need to practice.
Effort Affirmation (Correct No-effort)	That was too easy for you. Let's hope the next one is more challenging so that we can learn something.
Effort Affirmation (Correct Effort)	Good job! See how taking your time to work through these questions can make you get the right answer?
Strategic (Incorrect)	Are we using a correct strategy to solve this? What are the different steps we have to carry out to solve this one?
Strategic (Correct)	We are making progress. Can you think of what we have learned in the last 5 problems?

Attribution Interventions. Attribution theory proposes that students' motivation to learn is directly rooted in their beliefs about why they succeed or fail at tasks [21]. If students can be taught to alter these beliefs, for instance to understand that failure is the result of a lack of effort instead of a lack of ability, then their motivation to learn and learning outcomes can be significantly improved [22]. For example:

- G*eneral attribution* messages encourage students to reflect about myths and math learning in general;
- E*ffort attribution* messages reinforce that effort is a necessary by-product of learning, and are specially tailored to situations where students are investing effort but are struggling;
- *No-effort attribution* messages are more emphatic than the ones just mentioned; they are designed to help students realize that effort is necessary to learn, and generated when students are not investing effort;
- *Incorrect attribution* interventions are generated to motivate students after they provide an incorrect response, by re-formulating how they perceive errors.

Fig. 3. Jane, the female affective learning companion, and Jake, the male affective learning companion.

Effort-Affirmation Interventions. In contrast to the effort-attribution messages described above, which aim to change students' attitude towards effort during problem solving and are generated before the student actually starts problem solving, the *effort-affirmation* interventions acknowledge effort after students obtain a correct solution (see Table 2 for examples). These interventions include:

- *Correct no-effort interventions* are generated after a student invests no effort but obtains a correct solution, to make students realize that praise is not appropriate;
- *Correct-effort affirmations* are generated after a student both invests effort and obtains the correct solution, to acknowledge the student's effort.

Strategic Interventions. The final type of intervention we embedded into Wayang focuses on meta-cognitive strategies, with the goal of both making students more effective problem solvers and motivating them for learning in general.

- *Incorrect strategic* messages are generated when students are not succeeding at problem solving, to motivate them to change their general problem-solving strategy, i.e., think about why they are not succeeding
- C*orrect strategic* messages are generated when students are succeeding at problem solving, to encourage them to evaluate their progress.

5. The User Study

The user study was designed to quantitatively evaluate the impact of learning companions on affective and cognitive outcomes for all students. Of the 108 ninth- and tenth-grade students, two thirds (72 students) received a learning companion of a random gender, and one third (36 students) did not receive a learning companion. We obtained complete data, surveys and posttest, for about 95 students.

At the beginning of the study, students received a math pretest and a survey that assessed general attitudes towards math, described in Section 2. The following day and for the next three days, students used Wayang instead of their regular math class. Every five minutes as well as after completing a problem, students were asked to provide information on one of the four target emotions (e.g. "How frustrated do you feel?"), see Fig. 2. At the start of each session, the learning companions introduced themselves; when students needed help during problem solving, the companions reminded students about the "help button" that provided multimedia-based support in the form of animations with sound. Characters spoke aloud the messages described in the previous section, occasionally at the beginning of a new problem and/or after the student submitted a response to a problem. Students in the control group (no-LC) had access to the same cognitive support (e.g., hints, problems read aloud), but no companions and no affective support.

After students used Wayang for three days, they took a math posttest, and answered the same questionnaire as taken prior to using the tutor. In addition, the questionnaire included five questions about student perception of the tutor (*Did you learn? Was the tutor concerned about your learning? Helpful?*). We also logged student behavior with the tutor, such as success at problem solving, gaming (abuse of hints), use of tools, and help. Students' emotions within the tutor were recorded, as well as when students muted the characters (mute button), and whether they abused help by rapidly reaching the bottom-out hint, or quick guessed (i.e., rapidly selected options until they hit the correct answer).

Table 3. General Post-Tutor Outcomes: Main and interaction effects for Affective and Cognitive Outcomes. Key: H-A — High-Achieving students; L-A – Low-Achieving students; LC — Learning Companions; ∅ — No significant difference across conditions; ∅ MathAbility — No significant MathAbility effect or MathAbilityxLC interaction effect.

	Overall Effect	**Differential Effect (High vs. Low)**
Learning	Students learned in all conditions (paired samples t-test, ***t(99) = 2.4, p = .019**), but no significant effect for LC	L-A students *improved* more than H-A in all conditions ***F(99,1) = 5.3, p = 0.02**
Perceptions of Wayang	∅	When LCs are absent, H-A students *perceive* Wayang better than L-A. LCxMathAbility ****F(96,1)=6.84, p = 0.01**
Liking of Mathematics	Students receiving Jane demonstrated *higher math liking*. ***F(93,2) = 3.7, p = 0.03**	∅ MathAbility
Math Ability Self-concept	Students receiving Jane showed higher posttest *self-concept*. ***F(94,2) =3.6, p = 0.03**	When LCs are absent, H-A students had *higher increase* in *self-concept* than L-A. LCxMathAbility: **⁺F(94,3) = 2.3, p = .08**

5. Results, Discussion and Conclusion

We carried out an Analyses of Covariance (ANCOVA) for each affective and behavioral dependent variable (post-tutor and within tutor) as shown in Tables 3-4. In particular, Table 3 shows the results for general post-tutor outcomes, while Table 4 presents the results for affect-related and other variables measured within the tutor. As far as emotions, we include findings both on students' self-reported emotions within the tutor, and post test differences in survey responses (note that in Table 1, we reported how students were feeling *before* they interacted with Wayang, while Tables 3 and 4 look at how interaction with Wayang influenced these feelings). Our covariates consisted of the corresponding pretest baseline variable (e.g., we accounted for students' pretest baseline confidence when analyzing confidence while using the tutor or afterwards). Independent variables corresponded to *condition*, specifically learning companion (LC) present vs. absent and LC type (Female (Jane) vs. Male (Jake) vs. no-LC). We analyzed both main effects and interactions for achievement level (MathAbility) and conditions over all student data (see second and last columns of Tables 3 and 4). In addition, because of the special affective needs of low-achieving students, we repeated the ANCOVAs for the low-achieving student population only, for a "targeted effect," Table 4 (third column).

Results showed that all students demonstrated math learning after working with Wayang, with low-achieving students learning more than high achieving students across all conditions (Table 3). Learning companions did not affect student learning directly, but successfully induced positive student behaviors that have been correlated to learning, specifically, students spent more time on hinted problems [16] (see "Productive behavior" row, Table 4). The beneficial effect of learning companions was mainly on affective outcomes, particularly on confidence (see "Confidence" row, Table 4). Low-achieving students who received learning companions improved their confidence while using the tutor and at posttest time more than students with no learning companions, while their counterparts in the no-LC condition tended to decrease their confidence (Figure 4).

Table 4. Emotions within and after using the tutor. Key: H-A—High Achieving; L-A –Low Achieving; ∅ —No significant difference across conditions; ∅**MathAbilityxLC** —No significant MathAbilityxLC interaction effect/MathAbility effect; **LC**–Learning Companions.

	Overall Effect	**<u>Targeted Effect</u> on Low Achieving Students**	**<u>Differential Effect</u> (High vs. Low)**
Frustration	Less *overall frustration* self-reported with Jane ****F(213,2) = 6.1, p = .003**	L-A students *have lower post-tutor frustration* in the LC condition than no-LC. **+F(58,1) = 3.4, p=.07**	When LCs are absent, L-A students have higher post-tutor *frustration than* H-A. LC x MathAbility **+F(93,3) = 2.4, p = .08**
Confidence	Higher overall confidence reported in the LC condition ***F(204,1)=5.3, p = .02**	L-A students in the LC condition *have higher confidence*. Within Tutor LC effect: ****F(108,1)= 7.3, p = .008** Post-tutor LC effect: ***F(56,1)= 3.8, p = .05** and	H-A students *have higher confidence* than L-A students (but esp. when companions are absent) MathAbility effect within: ***F(204,1)= 4.1, p = .05** MathAbility effect posttutor: ***F(91,1) = 5.8, p = .02**
Interest	Students in the *LC condition* have higher overall interest at posttest time. LC main effect: **+F(94,1) = 3.4, p = .07**	L-A students in the LC condition report marginally more post-tutor interest. LC main effect: **+F(58,1) = 2.7, p =.1**	L-A students report *more boredom* than H-A students across all conditions MathAbility effect **+F(219,1) = 2.9, p = .09**
Excitement	∅	∅	H-A students report *less excitement* when LCs are absent, no difference when LC is present. MathAbilityxLC within: ***F(200,1) = 5.2, p=.02**
<u>Productive behavior</u>: time in hint problems	∅	L-A students spend more time in hinted problems with LCs. **+F(67, 1) = 2.9, p = 0.095**	∅ MathAbility
<u>Gaming behavior:</u> Quick–guess, help abuse	∅	∅	L-A students quick-guess more than do H-A students MathAbility effect: ****F(109,1) = 5.9, p = 0.017** No MathAbilityxLC interaction effect

Learning companions had a positive impact for all students on some measures, e.g., all students receiving the female companion (Jane) improved math liking and self-concept of their math ability. This was not the case for the male learning companion (Jake), which was muted by students twice as much as Jane, making it too similar to the control version.

Some differential effects (last column Table 4) suggest that learning companions are essential for low-achieving students' affect. When LCs are present, low achieving students report positive affect nearly as much as do high-achieving students and it is only when learning companions are absent that a large gap exists between these student groups. This affective gap reduces when learning companions are present. This result is found for several outcome variables: self-concept, perceptions of learning, frustration, excitement.

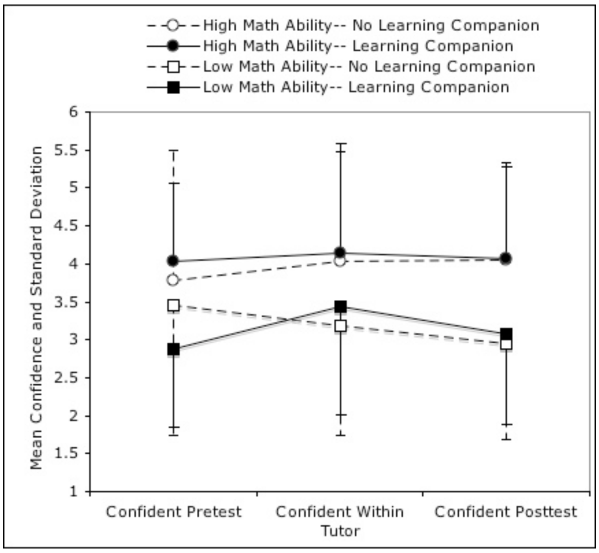

Fig. 4. Low and high achievement students' change in reported confidence during problem solving

However, learning companions did not manage to change some negative feelings and behaviors: low-achieving students did quick-guess more across all conditions than high achieving students; low achievement students reported less interest than high achieving in all conditions. We did see an increase in productive behaviors that lead to learning [16], low-achieving students spent more time in problems where help is requested (i.e. students pay more attention to hints).

General implications for tutors include the possibility of defining features and tool sets that support low-achieving students differentially from the rest. In future studies we will analyze separately the impact of companions on a large population of students with learning disabilities, compared to students without learning disabilities.

Acknowledgement. This research was funded by three grants: (1) US Dept of Ed (IES) *Using Intelligent Tutoring and Universal Design To Customize The Mathematics Curriculum* to Woolf (PI), Arroyo and Maloy; (2) NSF #0734060, HRD GSE/RES, *What Kind of Math Software Works For Girls?,* Arroyo (PI), Royer and Woolf; and (3) NSF #0705554, IIS/HCC *Affective Learning Companions: Modeling and Supporting Emotion During Teaching*, Woolf, Burleson (PIs) Arroyo, Barto, and Fisher. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

References

1. Rosiek, J (2003) Emotional Scaffolding: An Exploration of Teacher Knowledge at the Intersection of Student Emotion and the Subject Matter. Journal of Teacher Education, Vol. 54, No. 5, 399-412
2. D'Mello, S. & Graesser, A. (2007). *Mind and Body: Dialogue and Posture for Affect Detection in Learning Environments*. Frontiers in Artificial Intelligence and Applications.
3. McQuiggan, S. & Lester, J. (2006). *Diagnosing Self-Efficacy in Intelligent Tutoring Systems: An Empirical Study*. ITS 2006, Jhongli, Taiwan
4. Graesser, A., Chipman, P., King, B., McDaniel, B., and D'Mello, S (2007). Emotions and Learning with AutoTutor. Proceedings AIED 2007, pp 569-571. IOS Press.
5. Arroyo, I., Cooper, D.G., Burleson, W., Woolf, B.P., Muldner, K., Christopherson, R. (2009) Emotion Sensors Go To School. AIED 2009. Pp. 17-24. IOS Press.
6. Burleson, W. (2006). *Affective Learning Companions: Strategies for Empathetic Agents with Real-Time Multimodal Affective Sensing to Foster Meta-Cognitive Approaches to Learning, Motivation, and Perseverance*. MIT PhD dissertation.
7. Moreno, R., Mayer, R., and Lester, J. (2000) Life-Like Pedagogical Agents in Constructivist Multimedia Environments: Cognitive Consequences of Their Interaction. Proceedings of ED-MEDIA 2000, pp. 741-746.
8. Robison, J., McQuiggan, S., and Lester, J., (2009) Evaluating the Consequences of Affective Feedback in Intelligent Tutoring Systems. In *Proc. Inter. Conf. on Affective Computing & Intelligent Interaction*, Amsterdam, pp. 37-42, 2009.
9. Lester, J.C., Towns, S.G., FitzGerald, P.J. (1999) Achieving Affective Impact: Visual Emotive Communication in Lifelike Pedagogical Agents. *International Journal of Artificial Intelligence in Education*, 10(3-4), 278-291, 1999.
10. National Center for Educational Statistics, 2009. Digest of Educational Statistics, Chapter 2 http://nces.ed.gov/fastfacts/display.asp?id=64
11. Mass Gov (2006) Enrollment Data, Mass. Dept of Elementary and Secondary Education, http://www.doe.mass.edu/InfoServices/reports/enroll/?yr=sped0607
12. Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., Numtee, C. (2007). Cognitive Mechanisms Underlying Achievement Deficits in Children with Mathematical Learning Disability, *Child Development*, 78 (4): 1343-1359.
13. Geary, D.C., Hoard, M.K., & Hamson, C.O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology,* 74: 213–239.
14. Fletcher, J., Lyon, G., Fuchs, L., & Barnes, M. (2007). *Learning disabilities: From identification to intervention*. NY.
15. Olson, L., (2005). *State Test Programs Mushroom as NCLB Mandate Kicks In. Education Week*, November 30, 2005.
16. Arroyo, I.; Woolf, B.P. (2005). Inferring learning and attitudes from a Bayesian Network of log file data. AIED 2005: 33-40. IOS Press.
17. Arroyo, I., Beal, C., Murray, T., Walles, R. and Woolf, B. (2004). Web-based Intelligent Multimedia Tutoring for High Stakes Achievement Tests. ITS 2004, pp. 468–477.
18. Arroyo, I., Woolf, B., Royer, J. M., Tai M. (2009). Affective Gendered Learning Companions, *Inter. Conf. on Artificial Intelligence in Education*, England, IOS Press.
19. Cooper, D.G., Arroyo, I., Woolf, B.P. Muldner, K., Burleson, W., and Christopherson, R. (2009). Sensors Model Student Self Concept in the Classroom, International Conference on User Modeling and Adaptive Personalization, June 2009, Trento Italy.
20. Dweck, C. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance. *Improving academic achievement*. NY: Academic Press.
21. Weiner, Bernard (1972) Attribution Theory, Achievement Motivation, and the Educational Process Review of Educational Research, 42, 2, 203-215
22. Robertson, J.S. (2000) Is Attribution Training a Worthwhile Classroom Intervention For K–12 Students with Learning Difficulties? Education Psychology Review, 12:1, pp 111-134