Gender Matters: The Impact of Animated Agents on Students' Affect, Behavior and Learning

Ivon Arroyo[1], Beverly P. Woolf[1], James Michael Royer[1], Minghui Tai[1], Kasia Muldner[2], Winslow Burleson[2], David Cooper[1]

[1] University of Massachusetts-Amherst
[2] University of Arizona
ivon@cs.umass.edu, bev@cs.umass.edu, Winslow.Burleson@asu.edu, Katarzyna.Muldner@asu.edu, minghui_tai@yahoo.com, dcooper@cs.umass.edu

Abstract. We report on the reactions of males and females to the presence of animated agents that provided emotional or motivational feedback. One hundred (100) high school students used agents embedded in an Intelligent Tutoring System for Mathematics and randomized controlled evaluations compared students with and without learning companions. Positive results indicate that affective pedagogical agents can improve affective outcomes of students in general and particularly so for female students, who reported being more frustrated and less confident while solving math problems prior to using the tutoring system. We discuss issues of incorporating gender into user models and of generating responses tailored to gender.

Keywords: Intelligent tutoring Systems; gender differences; pedagogical agents; emotions, motivation and affect; empirical evaluation; mathematics education; student emotion

1 Introduction

If tutoring systems are to interact naturally with students, they need to model emotion and social competencies in addition to cognitive skills. Specifically, user models should address students' affective baggage and their fluctuations in affective states. Progress has been made regarding modeling students' affect [5, 6, 7, 19, 21]. However, little research has been done regarding how digital learning environments should respond to students' affect and how individual differences between students impact this process (for recent research see [2, 4, 19]).

Animated pedagogical characters have the potential to support students not only to learn more but also feel better about their learning experience, by engaging students through social interactions and tailoring of instructional content, e.g., a curriculum. However, to date pedagogical agents have mostly been focused on the cognitive rather than affective aspects. Some efforts have been made to create affective agents [7], but evaluations of their impact is still preliminary.

Here, we report on an evaluation of pedagogical agents in real school settings, with about 100 students of a public high school in Massachusetts, our results, and discuss the implications of our findings the for user modeling community. One of our key findings is evidence that gender has a key impact within the context of tutoring systems for mathematics. The next section describes research on gender differences and our initial results, followed by a description of the intelligent mathematics tutor that was used as a testbed. Sections 4 and 5 describe the emotional support provided by the tutor and the research study. Section 6 describes the benefits from this emotional and behavioral support and Section 7 provides a discussion of the value of including gender in a user model, especially when emotion of the user is involved.

2 Who needs affective support while learning mathematics?

Past research suggests that female students might have higher affective needs in certain disciplines; for example, in early adolescence gender differences exist in mathematics self concept (a student's belief about their ability to learn mathematics) and mathematics utility (the student's belief that mathematics is important and valuable to learn). Research shows that girls have less liking for math, more negative emotions and more self-derogating attributions about their math performance [9][10] Royer&Garofoli, 05). The literature suggests that while all students decrease their interest and increase their perception of difficulty of mathematics during high school, it is females and minorities that develop more negative feelings towards mathematics. It is believed that this poor affective relationship to the subject is one reason why females do not choose advanced math classes and later science careers in college [11], as compared to males who maintain a more positive relationship to the subject. Thus, helping girls in particular to foster a positive affective relationship to mathematics is highly relevant.

Our first goal was to verify the above hypothesis for our student population, which included 108 students from High School 1 and High School 2, two public high schools in MA --50% females and 15% with some documented learning disability. We focused our analysis on the affective needs of female students by analyzing mean differences between them and male students over a battery of affective pretest questions given to students before tutoring. These questions covered general attitudes towards mathematics, such as likes/dislikes of the subject, how important is mathematics, and specific questions about how students feel (e.g. anxiety/confidence, frustration, boredom, excitement) when they solve math problems.

In general, our results showed that students do not like math much – they do not find math very interesting or exciting (they score lower than neutral in our 6-likert scale assessments). Our initial pretest results suggested that both female students and low achieving students (lower than median math pretest score) have higher affective needs, as compared to other students [4]. Low achieving students disliked math more, valued it less, had worse perception of their math ability, and reported feeling worse when solving math problems than did high achieving students. The *math ability* of female students was similar to that of males according to our math pretest scores (Females M=.47 SD=.22; Males M=.50 SD=.19), but females reported feeling worse when solving math problems, as compared to males. Female students reported feeling significantly *less confidence* than did male students (Females M=3.1 SD=1.4; Males M=3.8 SD=1.2; $F(107,1)=7.8$, $p=.006$) and more *frustrated* (Females: M=3.6 SD=1.31; Males: M=3.1 SD=1.06; $F(108,1)=4.8$, $p=.03$) when solving math problems, before using the tutoring system.

The group with *worse confidence* and higher frustration values were low achieving females, who reported significantly worse feelings than low achieving males (e.g. for confidence, low achieving females M=2.7 SD=1.3; low achieving males M=3.4 SD=1.2, $F(63,1)=5.1$, $p=.03$). In fact, high achieving females have very similar feelings to low achieving males (e.g. for confidence, high achieving females M=3.6, SD=1.1) while high achieving males feel better than high achieving females (for confidence, high achieving males M=4.4, SD=1.4).

The above analysis confirms that our population for this study has the documented effect of females feeling worse about mathematics. In fact, Figure 1 shows this effect for a much larger population than the one covered for the data reported here, obtained for a second study across two different high schools (most subjects who were part of this study came from High School 2 and some from High School 1). It also shows that these negative feelings are not yet developed in 230 middle school students from School 1 (which contained both a middle and high school) just as the literature suggests. It seems that scaffolding female students' affect is an important target by itself: not only is female perception of mathematics worse than males but it is also unjustified, as it doesn't match their true mathematical abilities.

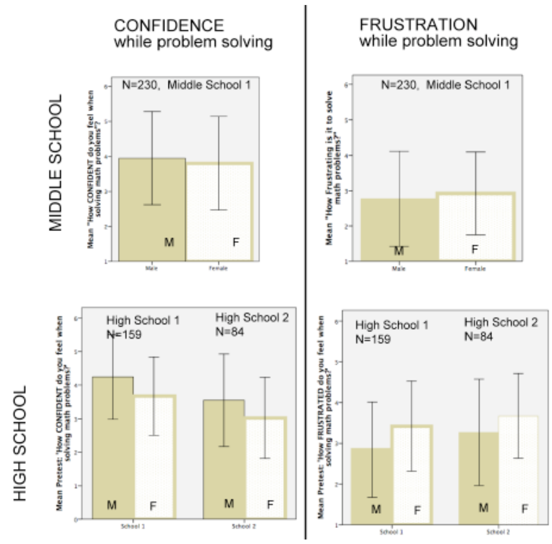

Figure 1: Results for a pre-tutor survey in two public schools: Girls develop negative feelings for mathematics, including *decreased confidence* (left) and *increased frustration* (right), between middle and high school.

3 The Testbed Tutoring System: Wayang Outpost

As the test bed for our research, we rely on Wayang Outpost, an intelligent tutor that helps students solve geometry problems, of the type that commonly appear on standardized tests [1]. To answer problems in the Wayang interface, students choose a solution from a list of multiple choice options (typically four or five), see Figure 2). Wayang provides immediate feedback on students' entries by coloring them red or green in the interface. As students solve a problem, they can ask the tutor for hints that are displayed in a progression from general suggestions to bottom-out solution. In addition to this domain-based help, the tutor provides a wide range of affective support, see Table 1, delivered by learning companions or agents designed to act like peers who care about a student's progress and offer support and advice. Currently, we have implemented two companions, Jane and Jake, Figure 3, as we are exploring how the gender of the companion influences outcomes (e.g., learning, attitudes) [18].

The tutor has several user models, including one for the student's *cognitive* skills and a second for the student's *effort* and *affect*. The learning companions' interventions are tailored to each student's needs according to these models. A simple *effort* model is used to assess the degree of effort a student invests to develop a problem solution (compared to what is expected for the math problem, based on thousands of past student interactions with it), mainly based on time per action. 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 [17, 19]. We now describe the affective interventions that Wayang's learning companions deliver and in the following section describe the reaction of students to these interventions.

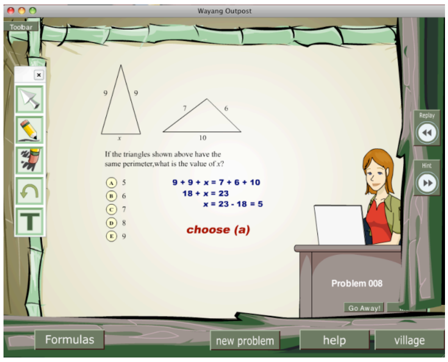

Figure 2. The Wayang Outpost Tutoring System

4 Providing affective support through spoken messages

Learning companions deliver approximately 50 different messages emphasizing the malleability of intelligence and the importance of effort and perseverance, Table 1. The messages also include meta-cognitive help related to effective strategies for solving mathematics 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 (Dweck, 1999). 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 the fact that the student developed a solution. They also offered praise to students who exerted effort while solving a problem, 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 merely expressed these non-verbal behaviors at random points, 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.

Type	Sample message
Attribution (General)	I found out that people have myths about math, think 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 from Wayang!
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?

Table 1. Sample behavior-based messages that characters speak to students

Attribution Interventions. Attribution theory proposes that students' motivation to learn is directly rooted in their beliefs of why they succeed or fail at tasks [15]. 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 [16]. We embedded several types of attribution training interventions into Wayang; these are spoken out when students face a new problem, and include (see examples in Table 1):

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

- G*eneral attribution* messages encourage students to reflect about myths, attitudes and math learning in general;
- E*ffort attribution* messages are designed to reinforce that effort is a necessary by-product of learning, and are specially tailored to situations where students are investing effort during problem solving but are struggling;

- N*o-effort attribution* messages are more emphatic than the ones just mentioned; designed to help students realize that effort is necessary to learn; generated when students are not investing effort during problem solving;
- I*ncorrect attribution* interventions are generated to motivate students after they provide an incorrect response, intended to think about errors positively.

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 1 for examples). The intention underlying the design of these messages is to both build a more realistic social bond between the companion and the student and to motivate the student. These interventions include:

- *Correct no-effort interventions* are generated after a student invests no effort but still obtains a correct solution, to make students realize the situation is not so worth of praise;
- *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 as a result becoming motivated for learning in general. The following strategic interventions are generated when the solution is either correct or not-correct (see Table 1 for examples):

- *Incorrect strategic* messages are generated when students are not succeeding at problem solving, and are designed to motivate students 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 students to evaluate their progress.

5 The User Study

The user study was designed to quantitatively analyze the benefit of learning companions on affective and cognitive outcomes for all students. The subjects included one hundred and eight (108) students from two high schools 1 (one low and the other high achieving) in the state of Massachusetts and involved 9th and 10th graders. Two thirds of the students were assigned to a learning companion of a random gender, and one third to the no learning companion condition. We obtained complete data (pre and posttest survey and math test) for a smaller subset of subjects.

Students took a mathematics pretest before starting, and completed a survey that assessed their general attitude towards mathematics[1]. Four questions asked about student feelings towards problem solving before they began to work with the tutor, including interest/boredom, frustration, confidence/anxiety, excitement (e.g. how frustrated do you get when solving math problems).

[1] The pre-test included 3 items for self-concept in math ability, e.g., students compared themselves to other students in their math ability and compared mathematics to other subjects; 3 items to address subjective mathematics liking/value).

For the next three days, students used the Wayang instead of their regular mathematics class. Approximately every five minutes, students were asked to provide information on one of the four target emotions (e.g. how frustrated do you feel?), see Figure 4. At the start of a student's interaction with Wayang, learning companions introduced themselves and when students needed help during problem solving, the companions reminded students about the "help button," which provided multimedia based support in the form of animations with sound. Characters spoke out the messages as described in the previous section, occasionally at the beginning of a new problem or after a correct or incorrect attempt to solve the problem

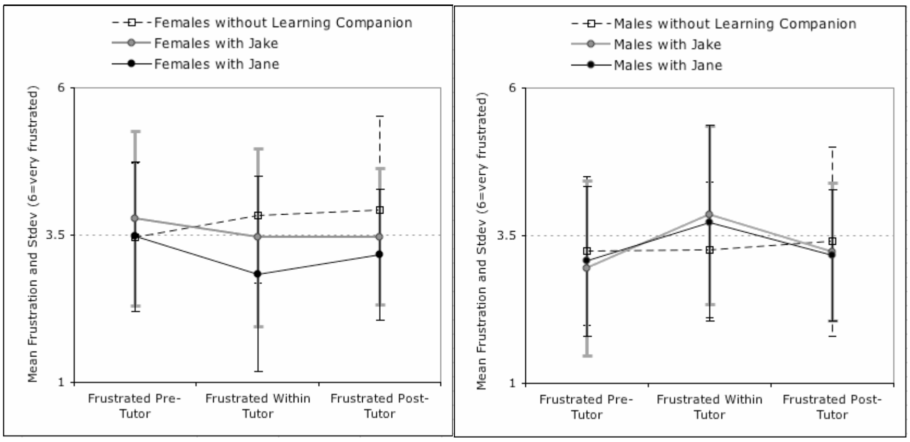

Figure 4. The frustration of male and female students before, within and after tutoring

After students used the tutoring module for three days, they took a mathematics post-test, and answered the same questionnaire they had received prior to using the tutor. In addition, the post-survey included five questions about the student's perceptions of the Wayang tutoring system (*Did you learn? Liked it? Helpful? Concerned? Friendly?*). Several student behaviors were logged, e.g., success at problem solving and use of tools and help. Students' self-report of their emotions within the tutor were logged, as well as students behavior, e.g., muting the characters (using a mute button), and whether they abused help or quick-guessed [22].

6 Results: who benefits from affective feedback?

We carried out Analyses of Covariance (ANCOVA) for each of the affective and behavioral dependent variables (post-tutor and within tutor) shown in Table 2. Covariates consisted of the corresponding pretest baseline variable (e.g. when analyzing within or posttest report of confidence towards problem-solving, we accounted for the pretest baseline confidence). Independent variables corresponded to *condition* (we tried both LC [Present/Absent], and Group [Jane/Jane/no-LC]). We analyzed both main effects and interactions for gender and condition over all student data (see second and last columns of Tables 2 and 3). In addition, because of the special affective needs of female students, we repeated the ANCOVAs for the female population only, described as a "targeted effect."

Table 2. General Post-Tutor Outcomes

	Overall Effects	**Targeted Effect**	**Differential Effect**
	All students (LC vs. no-LCMain Effect)	Females only (LC vs. no-LCMain Effect)	Females vs. Males (Condition x Gender Interaction effect)
Perceptions of Wayang	∅	***Higher perception of tutor*** **for females using companions**. **F(50,1)=7.5, p=.009	***Higher perception of tutor*** **for females using companions;** ***Higher perception of tutor*** **for males not using companions** **F(94,1)=10.5, p=.002
Math Liking	***Higher posttest*** **math *liking*** for **all students using female companion**. *F(93,2)=3.7, p=0.03	∅	∅Gender x LC
Math Ability Self-concept	***Higher improvement*** for **all students using female companion** *F(94,2)=3.6, p=0.03	∅	∅Gender x LC
Learning	**Significant improved learning for *all*** Paired Samples t-test *t(99)=2.4,p=.019 but no significant effect for LC	∅	∅Gender x LC

Key: ∅ -- No significant difference across conditions; ∅Gender x LC – No significant difference across conditions, no Gender x LC interaction effect; LC – Learning Companions.

Overall effects in Tables 2 and 3, second column, suggest a general advantage of learning companions (both Jane and Jake) for some affective outcomes. Table 3 shows that students reported significantly less frustration and more interest (less boredom) when learning companions were used compared to the no learning companion condition. At the same time, Table 2 shows that students receiving the female learning companion reported significantly higher self-concept and liking of math at posttest time. Students receiving Jane also reported higher confidence towards problem solving and in post-tutor surveys. One reason why Jake was at a disadvantage compared to Jane might be the fact that the male character was muted twice as much as was the female character. If students mute the characters, then the experimental condition turns out to be highly similar to the control condition (no learning companion) thus diminishing its effect. While significant results are limited to affective outcomes –learning companions did not impact learning--, we are impressed given the short exposure of students to the tutoring system.

Table 3. Emotions within and after using the tutor.

	Overall Effects	**Targeted Effect**	**Differential Effect**
Outcome	All students (LC vs. no-LCMain Effect)	Females only (LC vs. no-LCMain Effect)	Females vs. Males (Condition x Gender Interaction effect)
Interest	*Higher overall increase in interest* for students with LC. Post-tutor: $^{+}F(94,1)=3.4, p=.07$	∅	∅Gender x LC
Frustration	*Less overall frustration* reported with female companion $^{**}F(213,2)=6.1, p=.003$	*Less frustration* for females using female companion. Within tutor: $^{***}F(99,2)=8.2, p=.001$ After tutor: $^{+}F(49,1)=3.1, p=0.09$	∅Gender x LC
Confidence	*Higher overall confidence* for students using the learning companions. Within tutor: $^{*}F(204,1)=5.3, p=.02$	*More confidence* for *females in LC condition* Within tutor: $^{**}F(96,1)=5.6, p=.01$	∅Gender x LC
Excitement	∅	*More excitement* for females with LCs. After tutor: $^{+}F(53,1)=3.2, p=0.08$	*Less excitement* reported by females than males when LC is absent; Within tutor, GenderxLC: $^{*}F(200,1)=6.1, p=.02$ Post-tutor, GenderxLC: $^{*}F(67, 1)=5.3, p=.02$
Productivity: time in helped problems	∅	∅.	*Females spend more time* than males on "helped problems" in the LC condition; Within tutor, GenderxLC: $^{+}F(109,1)=2.78, p=0.09$
Gaming behavior: Help abuse and quick-guessing	∅	*Fewer quick guesses* by females in LC condition. Mean guesses per student: $^{**}F(55,1)=7.4, p=0.009$	Females *abuse help* marginally *less*, across all conditions. Gender: $^{+}F(110,1)=2.9, p=0.09$ Females make *fewer quick-guesses* with LC; males make *more quick guesses* with LC: GenderxLC: $^{**}F(109,1)=9.03, p=0.003$

Key: ∅ -- No significant difference across conditions; ∅Gender x LC – No significant difference across conditions, no Gender x LC interaction effect; LC – Learning Companions.

While learning companions afford affective advantages for all students, several significant effects in the ANCOVAs indicated a higher benefit of learning companions for female students. In the case of the emotional outcomes just mentioned (confidence and frustration in particular), the effects are stronger for females than for males (i.e. while all students improved confidence and reduced frustration, the third column of

Table 3 shows stronger significance for females alone). Last column of Table 3 also shows that females' confidence is improved but not confidence for males. It is important to note that these gender effects on emotions (within or after the tutor) are not due to females starting out feeling worse, as our analyses account for that baseline pretest emotion as a covariate.

Females especially perceived the learning experience with Wayang significantly better when learning companions were present, while the opposite happened for males, who actually reported better perceptions of learning with Wayang when learning companions were absent. Female students in the LC condition also had more productive behaviors in the tutor: they spent more time than males on "helped problems" compared to females in the no-LC condition; they "gamed" less when characters were present (a significant interaction effect revealed that the opposite happens for males).

7 Discussion: Does Gender Matter?

Past research has provided evidence that external features about students (e.g., verbal ability, reading level) may impact their performance. To date, much research has focused on cognitive factors, showing that general cognitive features are relatively permanent and can be assessed before students begin to use a tutor for the first time. For instance, Shute (1995) derived separate empirical student models depending on students IQ scores, as splitting the data this way provided a higher predictive value to determine students' state of knowledge than did a single student model. Students with differing IQ scores had different retention and acquisition parameters that impacted how much practice each needed. Another study showed that cognitive development differences in young children led to learning mathematics in different ways, i.e., from concrete manipulative vs. numeric hints, depending on their stage of cognitive development [2]. Different learning curves in each condition suggest that students should be modeled differently by an intelligent tutor, depending on what kind of help they receive.

Modeling gender is potentially powerful as it can enrich the predictive power of the student model and improve teaching power at a very low cost. The importance of including gender in a user model is not a mere hypothesis, but is based on extensive research on gender differences and learning at the K-12 level [13, 14]. Some research suggests that girls and boys have different approaches to problem solving [8, 12] and even that they should be taught differently [13]. While this literature involves gender differences in the classroom, we have found empirical evidence over the years that gender differences exist when males and females use tutoring systems at the K-12 level. In general, females are more "diligent" when using tutoring systems, showing behaviors that are more conducive to learning than those of male students (less gaming, spending more time on hints or accepting more help when offered). Also, females report better general attitudes while learning with the software, such as seriously trying to learn [3]. The findings of this paper are consistent with past research that suggests gender differences in response to affective feedback in particular --Burleson (2006) found that female students reduced frustration more than males when they were encouraged to reflect on their frustration, and concluded that gender should be further analyzed when providing affective support. This paper provides further evidence in support of a hypothesis that that affective support is gender dependent.

This research may ultimately lead to delicate recommendations about the type of support to provide for students. Should male students receive affective support at all? Should all females be provided with learning companions? These are harder questions to answer from these experimental results. While these results suggest that high school females will affectively benefit more than high school males when exposed to learning companions similar to ours within a math learning environment, we cannot conclude that males in general should not receive affective learning companions. In fact, another paper recently submitted [4] suggests that low achieving students (males and females) highly benefited from affective learning companions. It was only high achieving males who clearly did not benefit from affective learning companions, though our data set is not large enough to provide statistically significant results on the impact of learning companions for a combination of math ability and gender characteristics of students. We expect to conduct further studies with larger number of students in the future, in order to provide nuanced recommendations about gender and affective feedback in math ITS.

Acknowledgement

This research was funded by two NSF grants: (1) #0734060, HRD GSE/RES, *What kind of Math Software works for Girls? The effectiveness of motivational and cognitive interventions,* Arroyo, I. (PI) with Royer, J.M. and Woolf, B.P. and (2) #0705554, IIS/HCC *Affective Learning Companions: Modeling and supporting emotion during teaching*, Woolf, Burleson (PIs) with 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] Arroyo, I., Beal, C.R., Murray, T., Walles, R., Woolf, B.P. (2004) Web-Based Intelligent Multimedia Tutoring for High Stakes Achievement Tests. *International Conference on Intelligent Tutoring Systems:* 468-477

[2] Arroyo, I., Woolf, B.P., Beal, C.R. (2006). Addressing Cognitive Differences and Gender During Problem Solving. International Journal of Technology, Instruction, Cognition and Learning. Vol. 4, pp. 31-63.

[3] Arroyo, I., Woolf, B.P. (2005) Inferring learning and attitudes from a Bayesian Network of log file data. , *International Conference on Artificial Intelligence and Education*, IOS Press,: 33-40

[4] Arroyo, I., Woolf, B.P., Muldner, K., Burleson, W., Cooper, D., Razzaq, L., Dolan, R. (2010) Motivational Learning Companions Help Low Achieving and Learning Disability Students. Paper submitted to ITS 2010.

[5] Conati C. & Mclaren H. (2004). Evaluating A Probabilistic Model of Student Affect. *Proceedings of ITS 2004, 7th International Conference on Intelligent Tutoring Systems*, Lecture Notes in Computer Science. Springer Berlin / Heidelberg p. 55-66.

[6] D'Mello, S. K., Picard, R. W., and Graesser, A. C. (2007). Towards an Affect-Sensitive AutoTutor. S*pecial issue on Intelligent Educational Systems IEEE Intelligent Systems*.

[7] Graesser, A. C. , Chipman, P., King, B., McDaniel, B., and D'Mello, S (2007). Emotions and Learning with AutoTutor. *13th International Conference on Artificial Intelligence in Education*. R. Luckin, K. Koedinger, and J. Greer (Eds), pp 569-571. IOS Press.
[8] Fennema, E.; Carpenter, T.; Jacobs, V.; Franke, M.; Levi, L. (1998). "A Longitudinal Study of Gender Differences in Young Children's Mathematical Thinking." Educational Researcher, Vol. 27 (No. 5), 6-11.
[9] Eccles, J.S.; Wigfield, A.; Harold, R.D.; Blumenfeld, P. (1993). Age and gender differences in children's self and task perceptions during elementary school. Child development, 64, 830-847.
[10] Catsambis, S. (1994) The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high school. Sociology of Education, 67:3, pp. 199-215
[11] Catsambis, S. (2005). The gender gap in mathematics: Merely a step function? In A.M. Gallagher & J.C. Kaufman (Eds.), Gender differences in mathematics (pp. 220-245). Cambridge, UK: Cambridge University Press
[12] Carr, M. and Jessup, D. (1997). Gender Differences in First-Grade Mathematics Strategy Use: Social and Metacognitive Influences. Journal of Educational Psychology, 89:2, pp. 318-328.
[13] Sax, L. (2005) Why Gender Matters: What Parents and Teachers Need to Know about the Emerging Science of Sex Differences. Doubleday.
[14] Beal, C. (1994). Boys and girls: The development of gender roles. New York: McGraw-Hill.
[15] Weiner, B. (1986). An attributional theory of motivation and emotion. NY: Springer-Verlag.
[16] Robertson, J.S., (2000), Is Attribution Training a Worthwhile Classroom Intervention For K–12 Students with Learning Difficulties?, Educational Psychology Review, 12:1, pp 111-134.
[17] Arroyo, I., Cooper, D.G., Burleson, W., Woolf, B.P., Muldner, K., Christopherson, R. (2009) Emotion Sensors Go To School, *International Conference on Artificial Intelligence and Education*, IOS Press, pp. 17-24.
[18] Arroyo, I., Woolf, B.P., Royer, J.M., Tai, M. (2009) Affective Gendered Learning Companion, *International Conference on Artificial Intelligence and Education*, 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 Presentation, June, pp 30-41.
[20] 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 thesis.
[21] McQuiggan, S., Mott, B., Lester, J.: Modeling self efficacy in intelligent tutoring systems: An inductive approach. User Modeling and User-Adapted Interaction18(1) (2008) 81{123
[22] Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z. (2004) Off-Task Behavior in the Cognitive Tutor Classroom: When Students "Game The System". Proceedings of ACM CHI 2004: Computer-Human Interaction, 383-390