Mind the Gap: Improving Gender Equity in Game-Based Learning Environments with Learning Companions

Philip Sheridan Buffum[1], Kristy Elizabeth Boyer[1], Eric N. Wiebe[2], Bradford W. Mott[1], James C. Lester[1]

[1] Department of Computer Science, North Carolina State University
[2] Department of STEM Education, North Carolina State University
Raleigh, North Carolina, USA

{psbuffum, keboyer, wiebe, bwmott, lester}@ncsu.edu

Abstract. Game-based learning environments hold great promise for engaging learners. Yet game mechanics can initially pose barriers for students with less prior gaming experience. This paper examines game-based learning for a population of middle school learners in the US, where female students tend to have less gaming experience than male students. In a pilot study with an early version of ENGAGE, a game-based learning environment for middle school computer science education, female students reported higher initial frustration. To address this critical issue, we developed a prototype learning companion designed specifically to reduce frustration through the telling of autobiographical stories. In a pilot study of two 7th grade classrooms, female students responded especially positively to the learning companion, eliminating the gender gap in reported frustration. The results suggest that introducing learning companions can directly contribute to making the benefits of game-based learning equitable for all learners.

Keywords: Learning Companions, Game-Based Learning, Gender.

1 Introduction

Game-based learning environments can be highly engaging [1] and can effectively support student learning [2–4]. Research has shown that adding game-based features to existing educational software can increase motivation and mastery achievement if the given software supports long-term interactions [5]. However, among some learner populations, subgroups of students have less gaming experience than others, posing a barrier to the less experienced students' initial engagement within the game-based learning environment. For example, prior studies have found that male students tend to score higher on certain dimensions of engagement within game-based learning, perhaps due to their greater familiarity with gameplay mechanics [4]. Additionally, students who have low expectations upon beginning to interact with a game-based learning environment tend to learn less [6]. These findings indicate that great care must be taken to avoid disadvantaging students who enter with less prior gameplay

experience. Indeed, fine-grained analysis of many educational interventions, particularly for disciplines in STEM or computing, have found equity to be an issue on some level [7, 8].

Rather than dissuade the community from the merits of game-based learning environments, these findings should compel us to investigate ways to improve the equity of our systems. This paper focuses on a method to improve the *gender equity* of a game-based learning environment for middle school computer science. In this paper, we use the term "gender equity" specifically in regard to our efforts to make a system that does not disadvantage female students *due to its gameplay* (other dimensions of gender equity, such as efforts to ensure equal representation of female and male students playing the game, are crucial as well, but beyond the scope of this paper). Pilot data with the ENGAGE game-based learning environment revealed gender disparities in the early gameplay, likely due to female students feeling frustrated with gameplay mechanics. To counteract this, we looked to the success of animated pedagogical agents at improving attitudes among middle school students towards similar fields as computer science [9]. Specifically, we built a learning companion [10] designed to counteract the negative frustration that female students may encounter with gameplay mechanics early in gameplay. We hypothesized that the affective scaffolding provided by the learning companion would have a positive impact on female students, thus bringing more gender equity to the overall game-based learning environment.

2 Related Work

Animated pedagogical agents have long been found to have positive affective impact on young students, due to the well-documented *persona effect* [11]. Although some research has raised questions about whether the persona effect extends to a broader age range of learners (such as college students) [12], there is strong evidence as to the affective benefits of these agents on middle school students. More specifically, an animated pedagogical agent can be designed to take on any one of a number of different roles. *Intelligent tutors* generally operate as a knowledge expert, motivator, or mentor, all of which have demonstrated benefits [13]. *Teachable agents*, conversely, are designed to appear less knowledgeable than the student, requiring the student to learn for the sake of "teaching" the agent [14, 15]. Similarly, *learning companions* also play a non-authoritative role, providing a social peer alongside the student within an interactive learning environment [10].

Research on learning companions has featured gender as a focal point. The gender of the learning companion has been found to have an impact on students' attitudes and learning [16], as well as the relative likelihood of the agent receiving counterproductive, abusive behavior from the student [17]. Critically, learning companions have been found to have especially positive affective results for female students, such as improving female students' confidence in mathematics and reducing the gender gap in frustration in that field [18], as well as improving self-efficacy among female students [19]. There is strong evidence in support of tailoring the behavior of such agents based on the gender (as well as achievement level) of the learner [20–22]. In

addition to their gender-specific impact, learning companions have also been found to have especially great impacts for low-achieving students [23, 24].

3 ENGAGE Pilot Study: The No-Companion Condition

We have been developing a game-based learning environment, ENGAGE, to teach middle school students about computer science. In the game, students take on the role of computer scientists as they embark upon an adventure revolving around a socially relevant challenge, developing their computational thinking [25] skills along the way. The game has been piloted in multiple middle schools with diverse learner populations, and initial studies have provided promising results about the game's effectiveness for supporting students who enter with varying prior experiences and abilities [26]. The entire ENGAGE game-based learning experience includes gameplay and complementary out-of-game classroom activities. It has been integrated within a quarterly oceanography elective for middle school students that was initially offered in two public middle schools. Within this pilot study, data were collected from a total of 50 students at the two schools, in grades six through eight. The participants were 23 female and 27 male students. These students used a version of the ENGAGE game-based learning environment that did not include a learning companion.

While conducting classroom observations, the project team noticed (via qualitative interpretation of observational behavior) a potential trend of female students exhibiting frustration during early gameplay sessions, particularly with gameplay mechanics. As such, we looked to the survey data to confirm if such a trend existed. Table 1 shows the responses to a 5-point Likert scale item, in which students were asked to agree/disagree with the statement, "I felt frustrated while playing the game." (We will refer to this as *Frustration* for the sake of clarity). Students answered this question at the end of each session as part of the validated User Engagement Scale (UES) [27].

Differences were found *between female and male* students for Frustration on both of the first two gameplay sessions (over the academic term, students took an average of ten gameplay sessions to complete the entire game). In Session 1, female students reported an average Frustration of 2.91, compared to an average of 2.04 by their male classmates. In Table 1, this is the row labeled "Frustration – Session 1". We ran a One-Way ANOVA, and the results showed a statistically significant difference ($F(1, 48) = 7.53$, $p < .01$).

Table 1. No-Companion Condition: Frustration on Two Post-Surveys, on 5-point Likert Scale. (Higher Numbers Indicate More Frustration).

	Female (n = 23)	Male (n = 27)	Overall (n = 50)
Frustration – Session 1	2.91 (SD = 1.125)	2.04 (SD = 1.125)	2.44 (SD = 1.198)
Frustration – Session 2	3.48 (SD = 1.275)	2.63 (SD = 1.182)	3.02 (SD = 1.286)

Additionally, Frustration rose overall for all students from Session 1 to Session 2 (shown in the third column of Table 1), from an overall average of 2.44 to an overall average of 3.02. A Repeated Measures ANOVA found this also to be statistically

significant ($F(1, 49) = 9.35$, $p < .01$). Moreover, as one can see in the Table 1 row labeled "Frustration – Session 2", the gender gap persisted ($F(1, 48) = 5.96$, $p < .05$), with frustration levels for both genders rising as students confronted a gameplay challenge that, although not designed to be frustrating, proved to require skill with gameplay mechanics. A common problem in this 'Tri-Level Room' challenge was students accidentally falling in the water due to their struggles with navigating their avatar (Figure 1). The character in this scene is the student's avatar.

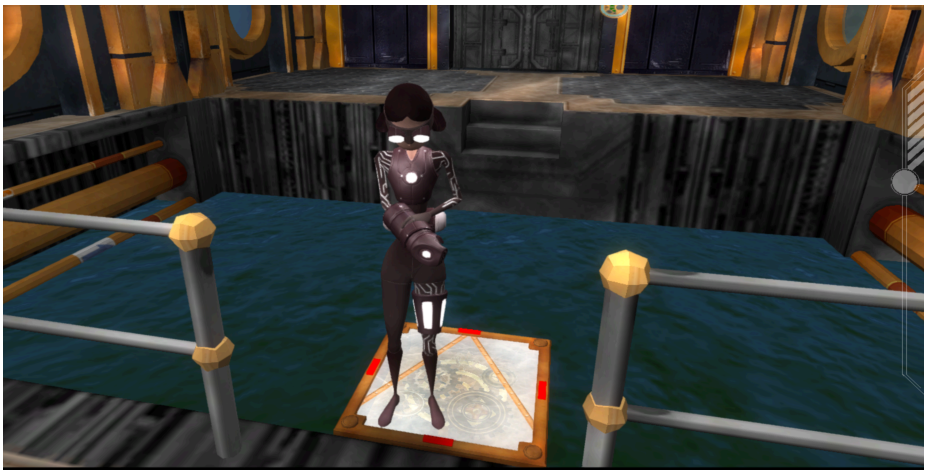

Fig. 1. Screenshot of Game at 'Tri-Level Room'.

4 The ENGAGE with Learning Companion Condition

Having noted this gender gap, we sought to design a timely intervention that could mitigate frustration in the gameplay for female students, as they overcame the initial obstacle of becoming familiar with gameplay mechanics. We consequently used Unity 3D to develop a prototype of an animated pedagogical agent, named Adriana. Adriana is a six-year-old girl whose sister lives in the underwater research station where the learner has been interacting in the game-based learning environment. In crafting Adriana's utterances, we aimed to design empathetic dialogue that would parallel the behaviors and resulting emotions that students might experience at this stage of the gameplay. We furthermore were guided by research on the value of agents' autobiographical stories and social conversation [28, 29], and particularly by research indicating that such agents can improve students' attitudes towards STEM fields [30]. Adriana tells the learner about her older sister, an expert computer scientist whom the player will meet later in the game. She also tells the learner about how she herself has visited the underwater research station and how she fell in the water many times. Adriana reassures the player that falling in the water in these challenging rooms is not abnormal, and that moving around the station will become easier with experience. Students respond to Adriana with typed, natural language replies. The interaction with

Adriana is designed to last approximately five minutes. Figure 2 shows an excerpt of interaction with Adriana.

Adriana:	"Um, falling in the water is not fun, especially when it happens over and over. Guess how many times I fell in. Do you know?"
Student:	"I dunno, maybe 2 or 3 times?"
Adriana:	"A million times. That's what [my sister] said – I can't count very well, but she said I fell in a million times. She says I'm clumsy. Do you think I'm clumsy?"
Student:	"Maybe a little."
Adriana:	"Haha! Anyways, don't feel bad if you fall in the water, too. It's only really a problem in the part of the station where you are. Other than that, how do you like it?"
Student:	"It is very nice."

Fig. 2. Excerpt of Student Interaction with Learning Companion

After developing a prototype of the learning companion, we conducted a study to test the agent's effectiveness, with particular interest in her ability to reduce frustration in female students during early gameplay. The study was conducted with two seventh-grade classes at an urban middle school. This was a different school than had been used in the No-Companion pilot study in order to ensure that the students had no prior exposure to the game-based learning environment. The school is in the same district as the prior schools and its demographics are similar. A total of 39 students participated in the study. We report results from the 28 students who completed the post-survey and who gave consent for participation in research. Of those 28 students (ages 12-13), 14 were female and 14 were male.[1]

Each participant attended two hour-long sessions, which were held on back-to-back days. In Session 1, students individually played the game up to a pre-defined stopping point. All students were able to reach this stopping point within fifty minutes of gameplay. Immediately upon reaching the stopping point, each student completed the Session 1 post-survey. We determined the stopping point based on the pilot data, which indicated that the next challenge, the 'Tri-Level Room', was particularly frustrating. Adriana's dialogue utterances were crafted to avert preemptively the frustration arising at this particular point in the game.

At the start of Session 2 on the second day, students first interacted with Adriana. Figure 3 shows a screen-capture from the beginning of an interaction. As with the rest of the gameplay, each student interacted with Adriana individually. We tested two versions of Adriana, one with dialogue and one with monologue, and participants were evenly distributed across the two versions. These interactions lasted an average of 5 minutes, with the shortest lasting just under one minute and the longest lasting almost 10 minutes. The agent's dialogue moves were pre-scripted, with each student receiving a similar progression of textual utterances from Adriana. As students finished their interactions with Adriana, they then proceeded with gameplay in ENGAGE, where they immediately encountered the 'Tri-Level Room'. Students continued play-

[1] Eleven of these 28 students interacted with a monologue version of Adriana. The monologue and dialogue data were aggregated into the Companion condition described in this paper.

ing the game until five minutes before the end of the session (resulting in approximately 40-50 minutes of gameplay for Session 2). Students then completed the Session 2 post-survey.

Fig. 3. Adriana the Learning Companion.

As in the No-Companion condition, the post-surveys for both Session 1 and Session 2 included the Frustration item, "I felt frustrated while playing the game." The Session 2 post-survey for this Companion condition also included four items about the student's affective response to Adriana, which we will refer to as *Reaction to Adriana*, as well as three demographic questions at the end. In the following section, we examine the results of Frustration and Reaction to Adriana.

5 Results

To analyze the success of Adriana the learning companion, we first examined the self-reported student Frustration. Table 2 summarizes the results of this survey item, broken down by gender and session.

Table 2. Companion Condition: Frustration on Two Post-Surveys, on 5-point Likert Scale.

	Female (n = 14)	Male (n = 14)	Overall (n = 28)
Frustration – Session 1	2.39 (SD = 0.87)	1.58 (SD = 0.67)	2.00 (SD = 0.87)
Frustration – Session 2	2.21 (SD = 1.25)	2.07 (SD = 1.27)	2.14 (SD = 1.24)

As with the No-Companion study, a gender gap was evident after Session 1 (before students interacted with the learning companion). In Table 2, this is the row labeled "Frustration – Session 1". Female students reported at this stage an average

Frustration of 2.39 on a 5-point Likert Scale, compared to 1.58 for their male classmates. A One-Way ANOVA found this to be statistically significant ($F(1,23) = 6.585$, $p < .05$). However, this gender gap was no longer observed after students interacted with Adriana and then played more of the game (including the part that had previously been found particularly frustrating for students). On the post-survey for Session 2 (in Table 2, this is the row labeled "Frustration – Session 2"), female students reported Frustration at an average of 2.21, statistically equivalent to the 2.07 average for male students ($F(1, 26) = 0.09$, $p = .77$). Moreover, there was *not* a significant increase in Frustration overall among all students from Session 1 to Session 2 (as shown in the third column in Table 2), which is noteworthy given the increased complexity of the Session 2 gameplay challenges. However, Frustration did rise for male students from Session 1 to Session 2 (which was also true for the No-Companion version of the game, as seen in Table 1).

In order to better understand these results, we examined the Reaction to Adriana items. There were four such items in which students were asked to agree/disagree with statements on a 5-point Likert scale. Table 3 displays the results for these items. Female students generally responded with higher ratings on all these items compared to their male classmates. While none of the differences is statistically significant, these results echo other findings on gender and learning companions [18] that have shown female students respond favorably to this type of pedagogical agent.

Table 3. Student Responses to Learning Companion, on 5-point Likert Scale.

	Female (n = 14)	Male (n = 14)
I enjoyed interacting with Adriana.	3.71 (SD = 1.14)	3.2 (SD = 1.03)
I would enjoy interacting with Adriana again in the future.	3.64 (SD = 1.28)	3.27 (SD = 1.12)
Interacting with Adriana helped me to enjoy playing the game.	3.79 (SD = 1.25)	3.27 (SD = 1.19)
Interacting with Adriana helped me to feel less frustrated while playing the game.	3.36 (SD = 1.08)	2.64 (SD = 1.01)

6 Discussion

Data from pilot studies of the ENGAGE game-based learning environment without a learning companion suggested that female students experienced more frustration. We developed a learning companion designed to mitigate this frustration. We hypothesized that in the version of the game with a learning companion, a gender gap in frustration would be observed after Session 1 (before the learning companion was present), but that this gender gap would no longer be observed after Session 2 when the learning companion was introduced. The results confirmed this hypothesis.

The learning companion Adriana effectively "leveled the playing field" for female students regarding frustration. Notably, female students in the Companion condition actually reported slightly less average frustration after Session 2 than Session 1, in contrast to the reverse finding in the No-Companion condition. While frustration still

rose for male students from Session 1 to Session 2, classroom observations suggest that this frustration stems from several other causes including the increased difficulty of the challenges in Session 2. Female students also still face these same alternative sources of frustration, but Adriana's intervention here mitigated the *additional* burden of being frustrated with gameplay mechanics. While not all frustration necessarily inhibits learning [31], these results suggest that the learning companion did improve gender equity in this game-based learning experience.

It is important to note that the two conditions described in this paper were not designed simultaneously as part of a controlled experiment. The larger ENGAGE study, which this paper refers to as the No-Companion condition, was conducted as an authentic classroom study in which students played the game in pairs. The study of the learning companion, referred to as the Companion condition, was conducted in a controlled manner, with students playing the game individually. An additional limitation is that there were two versions of the learning companion tested in the Companion condition, as part of an original study design to compare monologue with dialogue for the agent. However, there were no significant differences in student survey responses between the two versions of the learning companion, although interactions with the monologue version were several minutes shorter on average than those with the dialogue version. The Companion condition data were therefore aggregated together.

7 Conclusion and Future Work

Support for underrepresented populations, including underrepresentation based on gender, is of central concern in the design of interactive learning environments. For game-based learning environments, overcoming initial differences in gameplay experience is a particularly promising area for improving equity for learners. This paper has presented results demonstrating the success of a learning companion specifically crafted to mitigate the frustration that comes along with developing proficiency in gameplay mechanics. The learning companion in this study succeeded with female students, who were previously more frustrated than male students.

More broadly, this paper highlights the complex issue of gender equity in game-based learning environments. In designing the ENGAGE game, the research team conducted focus groups with diverse middle school students, created an array of avatars from which students could choose to represent themselves, and adopted best practices from the research community on broadening participation. Even with these careful steps taken, early versions of the game still revealed gender disparities in the level of frustration. The positive results with this prototype learning companion indicate that we need nuanced understanding of where individual game-based learning environments struggle with equity.

This naturally leads to a number of important directions for future research. One important area of investigation is how to create a learning companion with whom students will enjoy conversing many times over a period of time, which is vital for the success of game-based learning environments designed to support long-term interventions. Research on relational agents [32] will guide this future line of inquiry. Another

line of future research should examine how to support equity at a finer granularity. Categorizing students in binary fashion as "female" or "male" is a natural distinction, but ultimately we will need to consider all dimensions of a student's identity. Finally, a study of how learning companions can support collaborative game-based learning is an important direction, as related research on pedagogical agents that support collaboration within intelligent tutoring systems has shown promise [33]. With continued research and appropriate interventions, we as a research community can design game-based learning environments that equitably serve all students.

Acknowledgments. The authors wish to thank colleagues from the Center for Educational Informatics for their assistance. This work is supported in part by the National Science Foundation through Grants CNS-113897, CNS-1042468, and IIS-1409639. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the participants, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.

References

1. Sabourin JL, Lester JC (2014) Affect and engagement in game-based learning environments. IEEE Trans Affect Comput 5:45–56.
2. Johnson WL (2010) Serious use of a serious game for language learning. Int J Artif Intell Educ 20:175–195.
3. Kim JM, Hill RW, Technologies C, Lane HC, Forbell E, Core M, Marsella S, Pynadath D, Hill Jr. RW, Durlach PJ, Hart J (2009) BiLAT: A game-based environment for practicing negotiation in a cultural context. Int J Artif Intell Educ 19:289–308.
4. Rowe JP, Shores LR, Mott BW, Lester JC (2011) Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments. Int J Artif Intell Educ 21:115–133.
5. Jackson GT, McNamara DS (2013) Motivation and performance in a game-based intelligent tutoring system. J Educ Psychol 105:1036–1049.
6. Snow EL, Jackson GT, Varner LK, Mcnamara DS (2013) Expectations of Technology : A Factor to Consider in Game-Based Learning Environments. Proc. of AIED '13. pp 359–368.
7. Jenson J, de Castell S, Bryson M (2003) "Girl talk": gender, equity, and identity discourses in a school-based computer culture. Womens Stud Int Forum 26:561–573.
8. Shah N, Lewis C, Caires R (2014) Analyzing Equity in Collaborative Learning Situations: A Comparative Case Study in Elementary Computer Science. Int. Conf. Learn. Sci. Conf. pp 495–502.
9. Plant EA, Baylor AL, Doerr CE, Rosenberg-Kima RB (2009) Changing Middle-School Students' Attitudes and Performance Regarding Engineering With Computer-Based Social Models. Comput Educ 53:209–215.
10. Chou C-Y, Chan T-W, Lin C-J (2003) Redefining the Learning Companion: The Past, Present, and Future of Educational Agents. Comput Educ 40:255–269.
11. Lester JC, Converse SA, Kahler SE, Barlow ST, Stone BA, Bhogal RS (1997) The Persona Effect: Affective Impact of Animated Pedagogical Agents. Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI '97. pp 359–366
12. Miksatko J, Kipp KH, Kipp M (2010) The Persona Zero-Effect : Evaluating Virtual Character Benefits on a Learning Task with Repeated Interactions. Proc. 10th Annu. Int. Conf. Intell. Virtual Agents - IVA 2010. pp 475–481

13. Baylor AL, Kim Y (2003) Validating Pedagogical Agent Roles: Expert , Motivator , and Mentor. World Conf. Educ. Multimedia, Hypermedia Telecommun. pp 463–466.
14. Biswas G, Leelawong K, Schwartz D, Vye N (2005) Learning by Teaching: A New Agent Paradigm for Educational Software. Appl Artif Intell 19:363–392.
15. Chase CC, Chin DB, Oppezzo MA, Schwartz DL (2009) Teachable Agents and the Protégé Effect: Increasing the Effort Towards Learning. J Sci Educ Technol 18:334–352.
16. Arroyo I, Woolf BP, Royer JM, Tai M (2009) Affective Gendered Learning Companions. Proc. of AIED ’09. pp 41–48.
17. Silvervarg A, Raukola K, Haake M, Gulz A (2012) The Effect of Visual Gender on Abuse in Conversation with ECAs. Proc. 12th Annu. Int. Conf. Intell. Virtual Agents - IVA 2012. pp 153–160.
18. Arroyo I, Woolf BP, Cooper DG, Burleson W, Muldner K (2011) The Impact of Animated Pedagogical Agents on Girls’ and Boys' Emotions, Attitudes, Behaviors and Learning. 2011 IEEE 11th Int. Conf. Adv. Learn. Technol. IEEE, pp 506–510.
19. Kim Y, Wei Q, Xu B, Ko Y, Ilieva V (2007) MathGirls: Toward Developing Girls’ Positive Attitude and Self-Efficacy through Pedagogical Agents. Proc. of AIED ’07. pp 119–126.
20. Kim DY, Baylor AL (2014) Pedagogical Agents as Learning Companions: The Role of Agent Competency and Type of Interaction. Educ Technol Res Dev 54:223–243.
21. Burleson W, Picard R (2007) Evidence for Gender Specific Approaches to the Development of Emotionally Intelligent Learning Companions. IEEE Intell Syst J 22:62–69.
22. Arroyo I, Burleson W, Tai M, Muldner K, Woolf BP (2013) Gender differences in the use and benefit of advanced learning technologies for mathematics. J Educ Psychol 105:957–969.
23. Woolf BP, Arroyo I, Muldner K, Burleson W, Cooper DG, Dolan R, Christopherson RM (2010) The Effect of Motivational Learning Companions on Low Achieving Students and Students with Disabilities. Proc. of ITS ’10. pp 327–337.
24. Rader E, Echelbarger M, Cassell J (2011) Brick by brick: iterating interventions to bridge the achievement gap with virtual peers. Proc. 2011 Annu. Conf. Hum. factors Comput. Syst. - CHI ’11. pp 2971–2974.
25. Wing JM (2006) Computational Thinking. Commun ACM 49:33.
26. Frankosky MH, Wiebe EN, Buffum PS, Boyer KE (2015) Spatial Ability and Other Predictors of Gameplay Time: Understanding Barriers to Learning in Game-based Virtual Environments. AERA Appl. Res. Immersive Environ. Learn. SIG. To appear.
27. O’Brien HLO, Toms EG (2010) The Development and Evaluation of a Survey to Measure User Engagement. J Am Soc Inf Sci 61:50–69.
28. Bickmore T, Schulman D, Yin L (2009) Engagement vs . Deceit : Virtual Humans with Human Autobiographies. Proc. 9th Annu. Int. Conf. Intell. Virtual Agents - IVA 2009. pp 6–19.
29. Gulz A, Haake M, Silvervarg A (2011) Extending a teachable agent with a social conversation module: effects on student experiences and learning. Proc. of AIED ’11. pp 106–114.
30. Ogan A, Aleven V, Jones C, Kim J (2011) Persistent effects of social instructional dialog in a virtual learning environment. Proc. of AIED ’11. pp 238–246.
31. Baker RSJ d., D’Mello SK, Rodrigo MMT, Graesser AC (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. Int J Hum Comput Stud 68:223–241.
32. Bickmore TW, Picard RW (2005) Establishing and maintaining long-term human-computer relationships. ACM Trans Comput Interact 12:293–327.
33. Kumar R, Ai H, Beuth JL, Rosé CP (2010) Socially capable conversational tutors can be effective in collaborative learning situations. Proc. of ITS ’10. pp 156–164.