An Analysis of Attention to Student-Adaptive Hints in an Educational Game

Mary Muir, Cristina Conati[1]

[1]Department of Computer Science, University of British Columbia
{marymuir, conati}@cs.ubc.ca

Abstract. We present a user study to investigate which factors affect student attention to user-adaptive hints during interaction with an educational computer game. The game is Prime Climb, an educational game designed to provide individualized support for learning number factorization skills in the form of hints based on a model of student learning. We use eye-tracking data to capture user attention patterns on the game adaptive-hints, and present results on how these patterns are impacted by factors related to existing user knowledge, hint timing, and attitude toward getting help in general. We plan to leverage these results in the future for making hint delivery adaptive to these factors.

Keywords: Adaptive help, educational games, pedagogical agents, eye-tracking

1 Introduction

The ability of providing interventions that are adaptive to each student's specific needs is one of the distinguishing features of intelligent tutoring systems (ITS). One of the most widespread forms of adaptive interventions is to provide hints designed to gradually help students through specific educational activities when they have difficulties proceeding on their own [14]. Despite the wide adoption of adaptive hints, there is an increasing body of research showing their possible limitations, going from students *gaming the system*, i.e., using the hints to get quick answers from the ITS [see 7 for an overview], to *help avoidance*, i.e. students not using hints altogether [e.g., 8, 9]. In this paper, we are interested in investigating the latter issue. More specifically, we seek to gain a better understanding of which factors may affect a student's tendency to attend to adaptive hints that the student has not explicitly elicited.

This research has three main contributions to the ITS field. First, while previous work on help avoidance focused on capturing and responding to a student's tendency to avoid requesting hints [e.g., 8, 9], here we investigate how students react when the hints are provided unsolicited. A second contribution is that we look at attention to adaptive hints during interaction with an educational computer game (edu-game), whereas most previous work on student usage (or misusage) of hints has been in the context of more structured problem solving activities. Providing adaptive hints to support learning during game play is especially challenging because it requires a trade-off between fostering learning and maintaining engagement. We see the results we present in this paper as valuable information that can be leveraged to achieve this tradeoff. The third contribution of our work is that we use eye-tracking data to study user attention patterns to the adaptive-hints, an approach not previously investigated in hint-related research. In [13], we presented a preliminary qualitative analysis of eye-tracking data for two students playing Prime Climb, and edu-game for number factorization. In this paper, we extend that work by presenting a more extensive quantitative analysis based on data from 12 students.

After discussing related work, we describe Prime Climb in further detail. Next, we illustrate the user study we conducted for data collection. Finally, we discuss results related to how user attention patterns are impacted by factors related to user existing knowledge, hint timing, and attitude toward getting help in general. We also present preliminary results on how attention to hints affects subsequent game playing.

2 Related Work

Edu-games are seen as one of the most promising new forms of computer-based education; however, while there is ample evidence that they are highly engaging, there is less support for their educational effectiveness [e.g., 1, 2, 16]. User-adaptive edu-games are receiving increasing attention [e.g., 3, 4, 5, 15] as a way to improve edu-games effectiveness. However, most of the existing work has not been formally evaluated in terms of how adaptive edu-game components affect edu-game effectiveness. There has also been rising interest in using eye-tracking to gain insights on the cognitive and affective processes underlying a user's performance with an interactive system [e.g., 6, 10, 11, 12]. In this paper, we extend the use of gaze information to understand if/how users attend to an educational game's adaptive interventions. Adaptive incremental hints are commonly used in ITS, but their effectiveness is in question because of extensive evidence that students can misuse them. Two main categories of help misusage have been investigated so far in the context of ITS for problem solving. The first is *gaming the system*, i.e., repeatedly asking for help or entering wrong answers on purpose to get to bottom-out hints that explicitly tell a student how to perform a problem solving step and move on [7]. The second is *help avoidance*, i.e., not asking for help when needed [8]. Several models have been developed to detect in real-time, instances of gaming behavior and intervene to reduce this behavior [see 7 for an overview]. Aleven et al., [8] present a model that detects both gaming the system as well as help avoidance. In [9], this model is used to generate hints designed to improve students' help seeking behavior in addition to hints that help with the target problem solving activity. Not much work, however, has been done on understanding if/how students process adaptive hints that they have not elicited. In [9], the authors suggest that students often ignore these hints. A similar hypothesis was brought forward in [3], based on preliminary results on student attention to hints in Prime Climb, the game targeted in this paper. Those results were based on hint display time (duration of time a hint stays open on the screen) as a rough indication of attention. In [13], however, initial results based on the analysis of gaze data from two Prime Climb players suggested that students sometimes pay attention to hints. The results we present here confirm this finding and extend it by presenting an analysis of factors that impact attention.

3 The Prime Climb Game

In Prime Climb, students practice number factorization by pairing up to climb a series of mountains. Each mountain is divided into numbered hexagons (see Figure 1), and players must move to numbers that do not share common factors with their partner's number, otherwise they fall. To help students, Prime Climb includes the Magnifying Glass, a tool that allows players to view the factorization for any number on the mountain in the device at the top-right corner of the interface (see Figure 1). Prime Climb also provides individualized textual hints, both on demand and unsolicited. Unsolicited hints are provided in response to student moves and are designed to foster student learning during game playing by (i) helping students when they make wrong moves due to lack of factorization knowledge; (ii) eliciting reasoning in terms on number factorization when students make correct moves due to lucky guesses or playing based on game heuristics. Prime Climb relies on a probabilistic student model to decide when incorrect moves are due to a lack of factorization knowledge vs. distraction errors, and when good moves reflect knowledge vs. lucky guesses. The student model assesses the student's factorization skills for each number involved in game

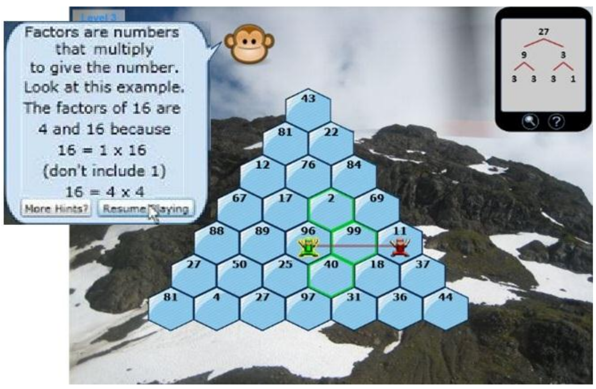

Fig. 1. The Prime Climb Interface

playing, based on the student's game actions [3]. Prime Climb gives hints at incremental levels of detail, if the student model predicts that the student doesn't know how to factorize one of the numbers involved in the performed move. The hint sequence includes a *tool* hint that encourages the student to use the magnifying glass to see relevant factorizations. If the student needs further help, Prime Climb gives *definition* hints designed to re-teach "what is a factor" via explanations and generic examples (e.g., see Figure 1). There are two different factorization definitions: "*Factors are numbers that divide evenly into the number*" and "*Factors are numbers that multiply to give the number*". The game alternates which definition to give first, and presents the second the next time it needs to provide a definition hint. The examples that accompany the definitions change for every hint, and are designed to help illustrate the given definitions while still leaving it to the student to find the factorization of the numbers relevant to the performed move. Finally, Prime Climb provides a *bottom-out* hint giving the factorization of the two numbers involved in the move (e.g., "*You fell because 84 and 99 share 3 as a common factor. 84 can be factorized as...*"). Students can access the next available hint by clicking on a button at the bottom of the current hint (See Figure 1). Otherwise, hints are given in progression as the student model calls for a new hint. A hint is displayed until the student selects to access the next hint or to resume playing (by clicking a second button available at the bottom of the hint). It should be noted that the Prime Climb *bottom-out* hints focus on making the student understand her previous move in terms of factorization knowledge; they never provide explicit information on how to move next. Thus, the Prime Climb hints are less conducive to a student gaming the system than bottom-out hints giving more explicit help [e.g. 7]. As a matter of fact, previous studies with Prime Climb show that students rarely ask for hints. Most of the hints the students see are unsolicited.

4 User Study on Attention to Hints

The study we ran to investigate students' attention to Prime Climb's adaptive hints relied on a Tobii T120 eye-tracker, a non-invasive desktop-based eye-tracker embedded in a 17" display that collects binocular eye-tracking data.

Twelve students (6 female) from grades 5 and 6 (six students for each grade) participated in the experiment. Participants first took a pre-test testing their ability to identify the factors of individual numbers and common factors between two numbers (16 numbers were tested overall). They then underwent a calibration phase with the Tobii eye-tracker. Next, they each played Prime Climb with an experimenter as a partner. The game was run on a Pentium 4, 3.2 GHz machine with 2GB of RAM, with the Tobii acting as the main display screen. Finally, participants took a post-test equivalent to the pre-test and completed a questionnaire on their game experience.

To analyze the attention behaviors of our study participants with respect to the received adaptive hints, we define an area of interest (*Hint* AOI) that covers the text of the hint message. We use two complementary eye-gaze metrics as measures of user attention to hints. The first is *total fixation time,* i.e., total time a student's gaze rested on the *Hint* AOI of each displayed hint. Total fixation time gives a measure of overall attention to hints, but does not provide detailed information on how a hint was actually processed (e.g., it cannot differentiate between a player who stares blankly at a hint vs. one who carefully reads each word). Furthermore, it is not ideal to compare attention to the different types of hints in Prime Climb because they have different lengths on average (15 words for *tool* hints; 17 words for *bottom-out* hints; 36 words for *definition* hints). Thus, our second chosen metric is the ratio of fixations per word (*fixations/word*), a measure that is independent of hint length and gives a sense of how carefully a student scans a hint's text.

5 Factors affecting attention to hints: Results

The study game sessions lasted 33 minutes on average (SD = 15). There was no improvement from pre to post-test performance, with participants scoring an average of 74% (SD = 31%) in the pre-test, an average of 72% (SD = 31%) on post-test and an average percentage learning gain of -0.02 (SD = 0.06). Consistent with previous Prime Climb studies, students rarely asked for help. One student asked for four hints, two students asked for hints twice, and two other students requested one hint. Prime Climb, however, generated unsolicited hints frequently: an average of 51 hints per player, (SD = 23), with an average frequency of 37 seconds (SD = 44). Thus, lack of system interventions can be ruled-out as a reason for lack of learning. If anything, it is possible that the hints happened too frequently, interfering with game playing and leading students to ignore them.

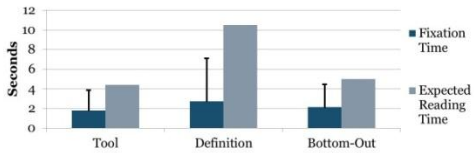

Fig. 2. Average Fixation Time for Prime Climb Hint Types

In order to investigate this idea further, we first compared average fixation time on each hint type with the *expected reading time* (calculated using the 3.4 words/second rate from [17]), which is the time it would take an average-speed reader to read the hint. Figure 2 shows that average fixation time is much shorter than *expected reading time* but the high standard deviation in all three measures shows a trend of selective attention. In the rest of this section, we investigate which factors influenced a student's decision to attend a hint or not. One obvious factor is whether the hints generated were justified, i.e., whether the probabilistic student model that drives hint generation is accurate in assessing a student's number factorization knowledge. Unfortunately we can only answer this question for the numbers tested in the post-test, which are about 10% of all the numbers covered in Prime Climb. The model *sensitivity* on post-test numbers (i.e., the proportion of actual positives which are correctly identified as such) is 89%, indicating that the model generally did not underestimate when a student knew a post-test number and thus it likely triggered justified hints on them. It should be noted, however, that for post-test numbers the student model is initialized with prior probabilities derived from test data from previous studies. For all the other numbers in Prime Climb, the model starts with generic prior probabilities of 0.5. Thus, the model's assessment of how student factorization knowledge on these numbers evolved during game play was likely to be less accurate than for post-test numbers, and may have generated unjustified hints.

Bearing this in mind, we looked at the following additional factors that may influence student attention to hints in our dataset. *Move Correctness* indicates whether the hint was generated in response to a correct or to an incorrect move. *Time of Hint* sets each hint to be in either the first or second half of a student's interaction with the game, defined by the median split over playing time. *Hint Type* reflects the three categories of Prime Climb hints: *Definition*, *Tool,* and *Bottom-out*. *Attitude* reflects student's general attitude towards receiving help when unable to proceed on a task, based on student answers to a related post-questionnaire item, rated using a Likert-scale from 1 to 5. We divided these responses into three categories: *Want help*, *Neutral*, and *Wanted no help*, based on whether the given rating was greater than, equal to, or less than 3 respectively. *Pre-test score* represents the student percentage score in the pre-test as an indication to student pre-existing factorization knowledge.

5.1 Factors that affect attention to hints measured by Total Fixation Time

We start our analysis looking at total fixation time on a displayed hint as a measure of attention. We ran a 2(*Time of Hint*) by 3(*Hint Type*) by 2 (*Move Correctness*) by 3 (*Attitude*) general linear model with *pre-test score* as a co-variant, and total fixation time as the dependent measure. We found the following interaction effects[1]:

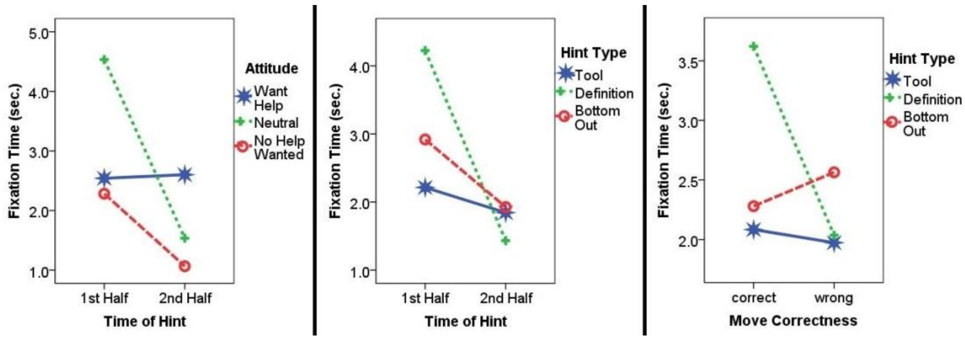

Fig. 3. Interaction effects between: (Left) *Time of Hint* and *Attitude*; (Middle) *Time of Hint* and *Hint Type*. (Right) *Move Correctness* and *Hint Type*.

- *Attitude* and *Time of Hint*. $F(2,447) = 5.566$, $p=0.004$, $\eta^2 =0.024$ (see Figure 2, Left). Fixation time for those with a neutral help attitude dropped from being the highest among the three groups in the first half of the game to being very low in the second half. For students who do not want help, fixation time is the lowest of the three groups in the first half of the game, and drops to even lower during the second half. Fixation time for those who wanted help did not change.
- *Time of Hint* and *Hint Type*, $F(2,447) = 5.963$, $p=0.003$, $\eta^2=0.026$. (see Figure 2, Middle). Fixation time drops for all hint types between the first and second half of the game. The drop, however, is statistically significant only for *definition* hints, suggesting that these hints became repetitive and were perceived as redundant despite the inclusion of varying examples that illustrate the definitions.
- *Hint Type* and *Move Correctness*, $F(2,447) = 3.435$, $p=0.033$, $\eta^2=0.015$. (see Figure 2, Right). Players had significantly higher fixation time on *definition* hints

[1] We also found main effects for both *Time of Hint* and *Attitude*, but we don't discuss them in detail because they are further qualified by the detected interactions

caused by correct moves than on those caused by incorrect moves[2]. There were no statistically significant differences between fixation times on correct vs. incorrect moves for the other two hint types. We find the result on *definition* hints somewhat surprising, because we would have expected hints following correct moves to be perceived as redundant and thus attended less than hints following incorrect moves. It is possible, however, that the very fact that hints after correct moves were unexpected attracted the student attention.

5.2 Factors that affect attention to hints measured by Fixations/word

To gain a better sense of how students looked at hints when they were displayed, we ran a general linear model with the same independent measures described above (*Time of Hint*, *Hint Type*, *Move Correctness*, *Attitude*, and *pre-test scores*) with fixations/word as the dependent measure. We found three main effects

- *Attitude*, $F(2,447) = 6.722$, $p=0.001$, $\eta^2=0.029$, Students who wanted no help had the lowest fixations/word (Avg. 0.25, *SD* = 0.30), significantly lower the other two groups. The difference between the help (Avg. 0.36, *SD* = 0.38) and neutral group (Avg. 0.31, *SD* = 0.28) is not significant, but the trend is in the direction of the help group having higher fixation/word than the neutral group.
- *Pre-test score*, $F(1,447) = 6.614$, $p=0.01$, $\eta^2=0.015$. Students with the lowest (below 65%) and highest (above 94%) scores had fewer fixations/word than students with intermediate scores. For high knowledge students, this effect is likely due to the hints not being justified. We can only speculate that, for low knowledge students, the effect may be due to a general lack of interest in learning from the game.
- *Hint Type*, $F(2,447) = 31.683$, $p<0.001$, $\eta^2=0.124$. *Definition* hints (Avg. 0.17, *SD* = 0.22) had a statistically significantly lower fixation/word than either *Tool* (Avg. 0.35, *SD* = 0.38) or *Bottom-out* hints (Avg. 0.34, *SD* = 0.32), possibly due to the fact that students tended to skip the actual definition part of the hints, which does not change, in order to get to the factorization examples at the bottom.

We also found two interaction effects, both involving *Move Correctness* (see Figure 4). The first interaction is with *Hint Type*, $F(2,447) = 11.141$, $p<0.001$, $\eta^2=0.013$. Fixations/word on *Bottom-out* hints drops significantly between those given after a correct move (Avg. 0.48, *SD* = 0.27) and those given after an incorrect move (Avg. 0.19, *SD* = 0.22). This result confirms the positive effect that *Move Correctness* seems to have on attention to hints found in the previous section for *definition* hints. Here, the effect possibly indicates that students are scanning *Bottom-out* hints for correct moves carefully in order to understand why they are receiving this detailed level of hint when they are moving well. The second interaction is with *Time of Hint*, $F(1,447)=3.922$, $p=0.048$, $\eta^2=0.009$ and shows that fixations/word drops significantly between hints for correct moves given in the first and the second half of the game,

[2] There is also a significant difference between fixation time on *definition* hints after correct moves and the other two type of hints after correct moves, but this difference is likely an effect of definition hints being longer, as we discussed in section 4.

suggesting that the aforementioned surprise effect of hints for correct moves fades as the game progresses.

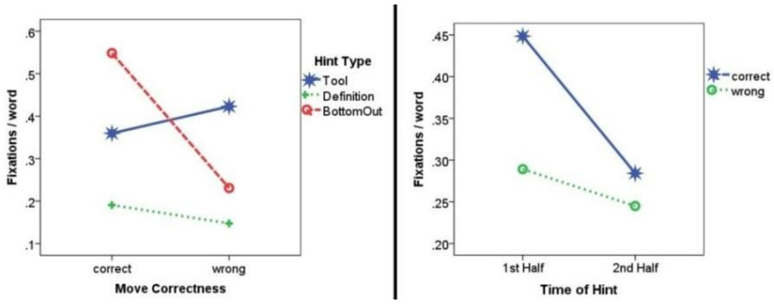

Fig. 4. Interaction effect between: (Left) *Time of Hint* and *Move Correctness*; (Right) *Move Correctness* and *Hint Type*

5.3 Factors that affect attention: Discussion.

All of the factors that we explored (*Time of Hint*, *Hint Type*, *Attitude*, *Move Correctness* and *Pre-test Scores*) affected to some extent attention to the Prime Climb hint, and the results of our analysis can be leveraged to improve attention to these hints. We found, for instance, that attention to hints decreases as the game proceeds, and the drop is highest for *definition* hints, suggesting that these hints are too repetitive and should be either varied or removed. If a student has an existing attitude toward help, this attitude generates consistent patterns of attention to hints throughout the game (low attention for those who do not want help, higher attention for those who do). This result suggests that general student attitude toward receiving help should be taken into account when generating adaptive hints, and strategies should be investigated to make hints appealing for those students who do not like receiving help. Similarly, strategies should be devised to make students with low knowledge (as assessed by the student model) look at the hints, since our results indicate that these students tend not to pay attention, although they are the ones who likely need hints the most. We also found that students with a neutral attitude toward help had much less consistent attention behavior than the students who wanted help and the students who did not. The neutral students showed quite high attention to hints in the first half of the interaction, but dropped almost to the lowest in the second half, confirming that the Prime Climb hints should be improved to remain informative and engaging as the game proceeds. In the next section, we show initial evidence that improving attention to hints as discussed here is a worthwhile endeavor because it can improve student interaction with the game.

6 Effect of attention to hints on game playing

In this section, we look at whether attention to hints impact students' performance with Prime Climb. In particular, we focus on the effect of attention to hints on correctness of the subsequent player's move. As our dependent variable, *Move Correctness After Hint*, is categorical (e.g., the move is either correct or incorrect), we use logistic regression to determine if *Fixation Time*, *Fixations per word* and *Hint Type* are significant predictors of *Move Correctness After Hints*[3].

Table 1. Logistic regression results for *Move Correctness After Hint*.

			95% CI for Odds Ratio		
	B (*SE*)	p	Lower	Odds Ratio	Upper
Fixations/word	0.98 (0.44)	0.03	1.12	2.68	6.39

Table 1 shows the results of running logistic regression on these data, indicating that *Fixations per word* is the only significant predictor of *Move Correctness After Hints*. The odds ratio greater than 1 indicates that, as fixations/word increases, the odds of correct moves also increases. This suggests that when the players read the hints more carefully, their next move is more likely to be correct. The results of the logistic regression also indicate that the type of hint student pay attention to does not impact move correctness. This finding is consistent with the fact that, in Prime Climb, *bottom-out* hints do not provide direct information on what to do next; they only explain how to evaluate the player's *previous* move in terms of number factorization, and this information cannot be directly transferred to the *next* move. Still, it appears that some form of transfer does happen when students pay attention to the hints, helping them make fewer errors on subsequent moves. This finding suggests that further investigation on how to increase student attention to hints is a worthwhile endeavor, because it can improve student performance with the game, and possibly help trigger student learning.

7 Conclusions and Future Work

In this paper, we presented a user study to investigate which factors affect student attention to user-adaptive hints during interaction with an educational computer game. This work contributes to existing research on student use and misuse of adaptive hints in ITS by looking at how students react to hints when they are provided unsolicited by the system, as opposed to explicitly requested by the student or obtained via gaming strategies. There are two additional aspects that are innovative in this work. The first is that we focus on adaptive hints provided by an edu-game, i.e., in a context in which it is especially challenging to provide didactic support because it can interfere with game playing. The second is that we use eye-tracking data to analyze student attention. We found that attention to hints is affected by a variety of factors related to user existing knowledge, hint timing/context and attitude toward getting help in general. The next step in this research will be to leverage these findings to improve the design and delivery of the Prime Climb hints. We also plan to extend the Prime Climb stu-

[3] The data points in our dataset are not independent, since they consist of sets of moves generated by the same students. Lack of independence can increase the risk of making a type 1 error due to overdispersion (i.e., ratio of the chi-square statistic to its degrees of freedom is greater than 1), but this is not an issue in our data set. (χ^2 = 6.41 df = 8)

dent model to use eye-tracking data in real-time for assessing if and how a student is attending to hints, and intervene to increase attention when necessary.

References

1. de Castell, S., Jenson, J.: Digital Games for Education: When Meanings Play. Intermedialities, 9, 45-54 (2007)
2. Van Eck, R.: Building Artificially Intelligent Learning Games. *Games and Simulations in Online Learning: Research and Development Frameworks*. D. Gibson, C. Aldrich, and M. Prensky, Editors, Information Science Pub. 271-307. (2007)
3. Conati C and Manske M.: Evaluating Adaptive Feedback in an Educational Computer Game, Proc. of IVA 2009, 9th International Conference on Intelligent Virtual Agents, Lecture Notes in Artificial Intelligence 5773. 146-158 (2009)
4. Peirce, N., Conlan, O. and Wade, V.: Adaptive Educational Games: Providing Non-invasive Personalised Learning Experiences. In Second IEEE International Conference on Digital Games and Intelligent Toys Based Education (DIGITEL 2008), Banff, Canada. 28-35 (2008)
5. Johnson, W. L.: Serious use for a serious game on language learning. In *Proc. of the 13th Int. Conf. on Artificial Intelligence in Education*, Los Angeles, USA. (2007)
6. Conati, C. and Merten, C.: Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation. Knowl.-Based Syst. 20(6): 557-574 (2007)
7. Baker, R., Corbett, A., Roll, I., Koedinger K.: Developing a generalizable detector of when students game the system. User Model. User-Adapt. Interact. 18(3). (2008)
8. Aleven, V., McLaren, B.M., Roll, I., and Koedinger, K.R.: 'Toward tutoring help seeking: Applying cognitive modeling to meta-cognitive skills'. *Intelligent Tutoring Systems 2004)*, Maceió, Brazil, pp. 227-239. (2004)
9. Roll, I., Aleven, V., McLaren, B., Ryu, E., Baker, R., Koedinger K.: The Help Tutor: Does Metacognitive Feedback Improve Students' Help-Seeking Actions, Skills and Learning? Intelligent Tutoring Systems 2006: (2006)
10. Bee, N., Wagner, J., André, E., Charles, F., Pizzi, D. and Cavazza, M.: "Interacting with a Gaze-Aware Virtual Character," *Workshop on Eye Gaze in Intelligent Human Machine Interaction (IUI 2010)*. (2010)
11. Prasov, Z. and Chai, J.: What's in a gaze? the role of eye-gaze in reference resolution in multimodal conversational interfaces. IUI 2008 (2008)
12. Muldner, K., Christopherson, R., Atkinson, R., Burleson, W.: Investigating the Utility of Eye-Tracking Information on Affect and Reasoning for User Modeling. UMAP 2009.
13. Muir M. and Conati C.: Understanding Student Attention to Adaptive Hints with Eye-tracking. UMAP 2011 Workshops. Springer LNCS 7138. 148-160 (2012).
14. Woolf, B. P.: Building intelligent interactive tutors: Student-Centered strategies for revolutionizing elearning. Morgan Kaufmann. (2008).
15. Easterday, M., Aleven, V., Scheines, R., Carver S.: Using Tutors to Improve Educational Games. AIED 2011: 63-71 (2011)
16. Linehan, C,. Kirman, B., Lawson, S. and Chan, G.: *Practical, Appropriate, Empirically-Validated Guidelines for Designing Educational Games*. CHI 2011, p.1979-1988 (2011).
17. Just, M. Carpenter, P. The Psychology of Reading and Language Comprehension. Boston. (1986)