

## CHAPTER 21 –Eye-Tracking for Student Modelling in Intelligent Tutoring Systems

Cristina Conati<sup>1</sup>, Vincent Alevan<sup>2</sup>, and Antonija Mitrovic<sup>3</sup>

<sup>1</sup>University of British Columbia, Vancouver, Canada, <sup>2</sup>Carnegie Mellon University, Pittsburgh, U.S.A, and

<sup>3</sup>University of Canterbury, Christchurch, NZ

### Introduction

---

In this chapter, we review research on leveraging eye-tracking information to improve the depth and accuracy of student modeling in ITSs. Eye-tracking has been extensively used both in psychology for understanding various aspects of human cognition, as well as in HCI for offline evaluation of interface design or as an alternative form of intended user input. In recent years, however, eye-tracking has also been investigated as a source of information on relevant users states and processes (e.g., attention, motivation, meta-cognitive activity) to inform the actions of an ITS. This chapter is an overview of some of the recent trends in this area. The overview is structured in two parts. In the first part, we describe existing research on investigating eye-tracking data as a direct source of information for student modeling and personalized instruction. In this part, we discuss efforts to model the learner at the behavioral, cognitive, meta-cognitive and affective level. The second part focuses on research that has used eye-tracking mainly for the offline analysis of how students attend to specific elements of an ITS interface, in order to understand relevant student behaviors and processes. Although this work is less directly related to using eye-tracking data in student modeling than the work described in the first part, the results of this research provide important insights on additional ways in which student models could leverage gaze-data in the future. We conclude the chapter with a discussion of these insights and related recommendations for GIFT design.

### Investigating Gaze Data as a Direct Source Of Information for Student Modeling

---

#### Leveraging Eye-Tracking Data to Capture and Adapt to Relevant Student Attention Behaviors

The work by Sibert et al. (2000) represents, to our knowledge, the first attempt to use gaze tracking for real-time student assessment. Sibert et al. (2000) describe GWGazer Reading Assistant, a system for automated reading remediation that tracks a student's reading patterns and provides support if these patterns indicate difficulties in reading a word. In particular, raw gaze data tracked with an unobtrusive, camera-based eye-tracker is parsed in real time to identify the word a student is currently reading, based on a *reading dwell* threshold that essentially defines the minimum amount of attention needed to be focusing on a word. A second threshold identifies delays in dwelling on a word that may indicate difficulty in reading it. When dwelling on a word exceeds this second threshold, the Reading Assistant pronounces the word for the student as an aid to reading it. Sibert et al. (2000) describe a preliminary informal evaluation of the system, in which eight children age 10–14 read a series of textual passages twice, with the help of the Reading Assistant. Results are presented in terms of changes in reading speed, accuracy, and number of prompts received from the first to the second reading of each passage, showing improvements on all measures. While these results do not provide any formal conclusions on the effectiveness of the Reading Assistant gaze-drive audio prompts, a qualitative post-questionnaire revealed that students liked the system and found it easy to use and unobtrusive. Thus, this work can be seen as an

## Design Recommendations for Intelligent Tutoring Systems - Volume 1: Learner Modeling

encouraging preliminary step in developing learning environments that leverage gaze information to provide personalized support to their users.

Anderson (2002) conducted an early experiment with a gaze-contingent ITS. The study involved a research version of the Pump Algebra Tutor (PAT), created specifically for use in eye-tracking studies (e.g., the interface elements were spaced out more widely than in the standard tutor). The purpose of the experiment was theoretical, namely, to provide an existence proof that paying attention to “fine-grained temporal detail of the student’s behavior” can have instructional leverage. More specifically, building on the work by Gluck and Anderson (2000), Anderson made the tutor respond to certain instructional opportunities that could be identified only through eye-tracking. As one example, if following an error message, the student did not read the error message (as revealed through eye-tracking) and did not correct the error within 10 seconds, the tutor would give a brief auditory message (“Read the help message”). As a second example, when – through eye-tracking – the tutor detected that the student used a problem-solving strategy that bypassed the instructional objectives, instead of using an intended problem-solving strategy, the tutor would give an auditory message suggesting that the student try the intended strategy (e.g., one in line with the instructional objectives). This happened in algebra problems in which students were given a problem statement describing a story context and were asked to (1) formulate an algebraic expression that captures the algebraic relations described in a story context and (2) use that expression to calculate specific quantities. On the second step (i.e., calculating quantities by applying the algebraic expression), eye-movement data revealed that students sometimes ignored the algebraic expression that they had created moments earlier and instead reasoned directly from the problem statement, a slower strategy that bypassed the objective of learning to work with algebraic expressions. When the student made an error on this type of step without fixating on the expression, the gaze-contingent tutor presented another auditory message: “Try using the formula to compute your answer.” With these gaze-contingent additions, the tutor helped students reach mastery 20% faster than the standard tutor, an impressive gain in efficiency. Further, eye-movement data revealed that students who worked with the gaze-contingent tutor attended to the algebraic expression more even before the gaze-contingent tutor would suggest they do, evidence that the gaze-contingent messages had the desired result.

Wang et al. (2006) used eye data to control interaction with the Empathic Software Agents (ESAs) for teaching biology. Eye-tracking data were used in two ways: as user input and also to provide information for adapting the behavior of the pedagogical agents available in the ESA to the student. Using the gaze input, it was possible for the student to choose a topic to study by simply looking at the appropriate area on the screen for a pre-defined time period. The student could also reply to a yes/no question by using the appropriate eye gesture (moving the eyes vertically for “yes” and horizontally for “no”). By analyzing the eye-tracking data and pupil dilation, the system inferred the student’s focus of attention and responded to it with affective behavior and/or feedback. When the student showed interest in a particular content by dwelling on it, the agent moved to the appropriate location on the screen and provided additional information. In addition, the agent also provided positive affective feedback to the student’s attentiveness by showing facial expressions conveying happiness. In contrast, if the student appeared to lose concentration (e.g., by looking away from the screen), the agent would say something to bring the student’s attention back to the screen and showed mild anger. The agents also displayed adaptive behaviors based on student’s states inferred primarily from student actions, e.g., providing feedback if a student made a mistake, or trying to engage the student if she appeared bored or disengaged because of lack of mouse or keyboard input. A small preliminary study with 10 participants revealed a beneficial effect of the adaptive agents on the students’ motivation and concentration. The participants reported that they were more attentive to additional information and explanations provided by the agent than to the other content available in the system. Although the study does not provide sufficient information to discriminate which role the gaze-adaptive components played in these evaluations, it shows that, overall, an agent relying on both gaze and action data to provide cognitive and affective feedback has good potential to enrich a student’s learning experience.

D’Mello et al. (2012) provide the only demonstration so far (to the best of our knowledge) that an ITS responding to student gaze in real time can improve student learning. (The study by Anderson [2002], described above, demonstrated that adding gaze-contingent responses to an ITS can lead to more efficient learning, but not that it leads to better learning outcomes.) The study involved Guru, a dialogue system with an on-screen tutor agent that engages the student in a tutorial dialogue on instructional material displayed on the screen. The system used eye-tracking to evaluate whether the student was paying attention, as captured using simple rules (basically, not looking at the tutor agent or the relevant instructional material was considered not paying attention). When the student did not pay attention, the tutor interjected (in speech) any of the following messages: “Please pay attention,” “I’m over here you know,” “You might want to focus on me for a change,” and “Snap out of it. Let’s keep going.” D’Mello and colleagues conducted a study comparing tutor versions with and without these gaze-responsive messages. The results were very interesting. First, there was substantial reorientation of gaze after the gaze messages, meaning that the tutor agent did succeed in directing students’ attention back to the tutor agent and the lecture. Further, students who worked with the gaze-reactive tutor did better on deep learning questions on the post test than students who worked with the version that was not gaze-reactive. In contrast, learning gains for assertion questions in the pre-post test, which tap into knowledge of surface level facts, were higher with the non-gaze-reactive tutor.

All the work described in this section leverage gaze information to capture momentous student attention patterns relevant to improving student interaction with the corresponding learning environments (i.e., patterns indicating reading difficulty in the GWGazer Reading Assistant, and patterns indicating attention or lack thereof in ESA and Guru). None of this work, however, uses the captured gaze information to make higher-level inferences regarding student’s states and processes. In the next section, we review work that takes this extra step, using gaze data to model students at the cognitive, meta-cognitive, and affective level.

### **Leveraging Eye-Tracking Data to Model Student Cognitive, Meta-Cognitive, and Affective States**

Conati and Merten (2007) use gaze data to improve the accuracy of a student model designed to enable provision of personalized support to learning mathematical functions via exploration of an interactive simulation (Adaptive Coach for Exploration [ACE]). Providing this adaptive support is challenging because it requires assessing the effectiveness of behaviors for which there is no formal definition of correctness. Conati and Merten (2007) tackled the challenge with a probabilistic model that assesses exploration effectiveness by integrating information on (1) user actions in ACE, (2) user’s knowledge and (3) whether users actually reason about (self-explain) their exploratory actions. Self-explanation – generating explanations to oneself to clarify instructional material – is a well-known meta-cognitive skill in cognitive science. This work is the first to consider and model self-explanation in the context of exploration-based learning. To assess whether a student is self-explaining the outcome of an exploratory action, the ACE’s student model combines information on the time the student spent on that action with gaze information. This gaze information relates to the occurrence of a simple gaze pattern defined a priori as being relevant for learning with this particular simulation: a gaze shift between two panels, one showing a function equation and one showing the related plot. The main exploratory action available in this simulation is to change either the equation or the plot, and see how the change affects the other component. Hence the definition of the aforementioned gaze shift as a relevant pattern to indicate self-explanation in ACE. A formal evaluation showed that the student model including eye-tracking information provides significantly better assessment of both a student’s self-explanation behavior during interaction with the simulation, as well as subsequent learning of the relevant mathematical concepts.

In the student model described above, Conati and Merten used gaze information related to the occurrence of a simple gaze pattern defined a priori as being relevant for learning with their target simulation. Kardan and Conati (2012) and Kardan and Conati (to appear) extend this work by looking at a much broader range of general eye-tracking features to capture student learning in the context of a different interactive simulation (IS). This is an important difference, because pre-defining gaze patterns that indicate learning in an IS may not always be easy or possible, due to the often unstructured and open-ended nature of the interaction that IS support. Furthermore, such pre-defined patterns are task specific, and may not directly transfer to a different IS. In contrast, the approach described in Kardan and Conati (2012) and Kardan and Conati (to appear) is more general and can be applied to a variety of ISs. It relies on giving to a classifier user model a broad range of standard eye-gaze features that are either task independent or based solely on identifying the main components of the target IS interface. Then, it is left to the classifier to identify patterns that are indicative of users' learning with that IS. An evaluation of this approach was performed on a data set encoding the gaze data of students working with the constraint satisfaction problem (CSP) applet, an IS designed to visualize the workings of the AC3 algorithm for constraint satisfaction on a variety of available sample problems. The CSP applet provides various functionalities that allow a student to explore the run-time behavior of AC3 at their own pace. The evaluation described by Kardan and Conati (2012) showed that a classifier using solely information on a student's overall attention patterns during a complete session with the CSP applet achieves an accuracy of 71% in distinguishing students who learned well from the CSP applet from students who did not (where learning was measured via a pre-test and post-test administered during the study). Furthermore, giving the classifier additional information on how students' attention patterns changed while solving two different problems of increasing difficulty further improved classification accuracy to 76%, with better balance in classifying each learner type (i.e., high learners vs. low learners, with class accuracy of 77% and 78%, respectively). In a follow-up study, Kardan and Conati (2013) showed that a student model for the CSP applet that combines information on both gaze data and interface actions outperforms models that rely on either gaze data or action data only. Kardan and Conati (to appear) also show that the action+gaze student model for the CSP applet reaches and stays above 85% accuracy in classifying a new user as a high versus low learner after seeing 22% of the overall interaction data (accuracy above 80% in each class), showing that the model can be used to trigger real-time interventions aimed at improving the experience of low learners with the CSP applet. Thus, Kardan and Conati's work provides further evidence of the value of gaze data for user modeling, especially for interactions in which it is hard to predefine a priori the learners' behaviors that should be detected as relevant or detrimental for learning.

Similar results were obtained by Bondareva et al. (2013), when using gaze data only to predict learning with a different type of educational environment, namely, a multi-agent ITS (known as Meta-Tutor), that scaffolds self-regulated learning (SRL) while students study science material (Azevedo et al., 2012). MetaTutor is an adaptive hypermedia learning environment, which includes 38 pages of text and diagrams, organized and accessible by an interactive table of contents. Text and diagrams are displayed separately in the two central panels of the interface. In addition to providing structured access to relevant content, MetaTutor also includes a variety of components designed to scaffold learners' use of SRL processes and their learning of a target science topic, e.g., the human circulatory system. Four pedagogical agents (PAs) provide spoken prompts and feedback on various SRL processes. For example, one PA assists the student in establishing two learning sub-goals related to the overall learning goal for the session. Other SLR processes supported by the PAs include taking notes, writing summaries of the viewed content, and evaluating one's current understanding via interactive quizzes.

The results in (Bondareva et al., 2013) show that, by leveraging gaze features similar to those used in Kardan and Conati (2013), a logistic regression classifier achieves 78% accuracy on predicting student learning with Meta-Tutor, after seeing all data from an interaction. Accuracy already reaches 72% accuracy after seeing 37% of the data. These results are especially important because, in combination with the results in Kardan and Conati (2013), they confirm the importance of gaze data as a predictor of

learning *across different types of learning environments* that can be leveraged for providing real-time personalized support to student learning.

Qu et al. (2005) leveraged gaze data to assess student motivation in the Virtual Factory, an ITS that teaches engineering skills (Johnson, Rickel, and Lester, 2000). They started from observations that human tutors use information about a learner's motivational states related to effort, confusion, and confidence during coaching. Based on these observations, Qu et al. (2005) enhanced an animated pedagogical agent with the ability to infer the same motivational factors about students. Information about the student's interface actions as well as gaze data tracking a student's focus and duration of attention were used as input for a dynamic Bayesian model, which inferred a learner's confidence, effort, and confusion during interaction the Virtual Factory. This student model was tested through a Wizard of Oz study during which students were interacting with a version of the Virtual Factory with the PA's interventions being directed by an experimenter. During the study, log data were collected, along with videos of the students' face and student retrospective self-reports on their motivational states during interaction. Two judges labeled replays of each session, synchronized with the videos of the students' face, for confusion, effort, and confidence (as Low, Medium, and High). The student model's predictions over the three factors were compared against both the judges' generated labels and the students' self-reports, showing very encouraging accuracies between 70% and 82%. Thus, this work provides initial evidence that gaze information can help assess student's affective states in addition to more strictly cognitive factors.

### **Off-Line Analysis of Gaze Data to Understand Relevant Student Behaviors and Processes.**

---

Seminal work by Gluck, Anderson, and Douglass (2000) demonstrated that eye-movement data of students working with an intelligent tutoring system contain information about students' cognitive processes that is not directly available from the regular stream of student-tutor interaction data (see also Anderson, 2002). By performing offline analysis of eye-tracking data obtained with a simplified version of the PAT algebra tutor (later named the Algebra Cognitive Tutor), these researchers were able to predict certain errors even before they happened. They also showed that eye-tracking data could quite reliably disambiguate domain-specific strategies even when they led to the same problem-solving steps. Finally, using eye-tracking, it became apparent that students did not attend to as many as 40% of the system's error feedback messages. Although the work by Gluck et al. (2000) did not actually demonstrate a method for updating a learner model based on eye-tracking, it is important for this survey because it clearly indicates the potential of gaze data as a rich source of information for student modeling, especially the strategy disambiguation work, in which inferences from eye-movement data to cognitive processes were made quite successfully, which often tends to be rather difficult step, fraught with uncertainty.

In relation to using gaze data to evaluate whether students attend to an ITS's adaptive interventions, Muir and Conati (2012) performed offline analysis of gaze data to investigate not only if, but also why students pay attention to adaptive hints generated by an educational game for math (Prime Climb). Prime Climb provides game activities to help students practice skills related to number factorizations, and includes a pedagogical agent that helps students learn from these activities by providing individualized hints. These hints are based on a student model that assesses whether students are learning during a session with Prime Climb, given their game actions. The hints are provided at incremental level of detail when the model predicts that student's knowledge of relevant factorization skills is low. The hints include (1) reminders to use available tools that can show how a number is factorized; (2) definitions of relevant factorization concepts, accompanied by illustrative examples; and (3) "bottom-out" hints that explicitly explain why a student action was correct or incorrect based on factorization knowledge.

Providing adaptive hints to support learning during game play is challenging because it requires a trade-off between fostering learning and maintaining engagement, thus this study aimed at investigating if there are factors that impact student attention to hints and that could be leveraged by a student model to make these hints more effective. Offline statistical analysis of the gaze data collected from 12 students (age 10–11) playing Prime Climb showed that attention to hints is significantly affected by the following factors: *time of hint* (i.e., whether a hint is given in the first or second half of a Prime Climb session), *hint type*, *attitude toward receiving help* (i.e., whether a student likes receiving help or prefers to do things without help), *game action correctness* and *pre-test scores* (i.e., how much factorization knowledge the student has before starting to play the game). Thus, this offline analysis indicates that capturing these factors and student attention to hints in the Prime Climb student model could help tailor hint presentation to a specific student. Muir and Conati (2012) also found that increased attention to hints was significantly correlated to increased correctness of the subsequent action, showing 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.

Eye-tracking has also been used to investigate the students' interaction and usage of OLMs. Since the mid-1990s, OLMs have attracted a lot of attention within the research community. Allowing the student to access an abstraction of the student model is beneficial in several ways. First, by opening the student model, ITSs become more user-friendly. Many projects have shown that students are capable of scrutinizing their models in order to explore the adaptive nature of the systems, and are interested in seeing the OLMs (Bull et al., 2005; Bull et al., 2007). Moreover, students can be actively involved in the modeling process via OLMs, as some systems allow students to challenge or even update their own student model. Finally, OLMs encourage students to think about their own knowledge, thus involving the student at the meta-cognitive level.

Eye-tracking has been used in several projects to investigate how students process the information in OLMs and evaluate the effectiveness of various types of OLMs. Bull, Cooke, and Mabbott (2007) investigate students' exploration of six different OLMs for the domain of C programming: a ranked list of concepts, a textual summary of the student model, a hierarchical lecture structure, a concept hierarchy, prerequisite relationships between concepts, and a concept map. In all views except the text summary, color is used to indicate knowledge level, with shades of green indicating correct understanding, yellow and white indicating low knowledge, and red indicating misconceptions. The participants were asked to interact with the OLMs, edit them, and/or persuade the system to change student models. The eye-tracking sessions lasted for 10 min, and students' preferences for various OLM views were collected via a user questionnaire. Participants generally found the OLMs useful, but had different preferences for which OLMs to use, and spent more time viewing misconceptions in their preferred views. Participants spent much more time examining their knowledge level (which promotes reflection) using the textual representation and ranked concept in comparison to the concept map and the prerequisites. The more complicated OLM views resulted in a broader spread of attention; for example, in the concept map participants focused less on their knowledge level but instead examined the map itself (i.e., they focused on the concepts for which there were insufficient data in their student models). Such more complicated OLMs require more effort from the student to gain an overview of the relationships between concepts.

Mathews et al. (2012) also used gaze data to analyze how students interpret OLMs in the context of EER-Tutor, a constraint-based ITS that teaches conceptual database design. The participants of the study were familiar with EER-Tutor, having used it previously in a database course. The participants viewed four different OLM views: concept tag cloud, kivi graph, concept hierarchy, and tree map. The goal of the study was to see whether the students understood the OLMs they were presented. The participants were asked three questions about each of the OLM views. For example, participants were asked how much the student (represented by a provided OLM view) had learnt about a particular concept. To answer questions, participants needed to examine the provided OLM. The eye-tracking data were collected in

addition to the answers provided by participants. The efficiency of an OLM view was calculated as the quotient of the participant's score (on the answers) and the product of the time spent viewing the OLM view and the number of fixations. A significant difference was found between the efficiencies of the four OLM views. Kiviat graphs and concept hierarchies were significantly easier to interpret in comparison to tag clouds and tree map according to the efficiency measure. Responses from the user questionnaire also identified tag clouds and tree maps as difficult to use to answer precise questions about knowledge levels. Participants were asked to rank the four OLM views by their preference: the highest ranked OLM was kiviatic graph, followed by tag clouds, concept hierarchy, and finally, tree maps. Participants commented that the kiviatic graph was best for an overall understanding of the student's knowledge, but that the concept hierarchy was valuable for more comprehensive understanding.

As the last chapter in this section, we report work indicating that an additional type of eye-based data, namely, pupillary response, can be leveraged for offline analysis of relevant student states during interaction with an ITS. Muldner et al. (2009) looked at the relationship between pupil dilation and relevant student affective and meta-cognitive states during interaction with EA-Coach, an ITS that helps students learn from analogical problem solving by scaffolding the relevant meta-cognitive skills of *self-explanation* and *analogical reasoning*. A study was conducted with 15 university students who verbalized their reasoning and affective states while interacting with the EA-Coach. The collected protocols were coded for meta-cognitive events (e.g., student utterances indicating *self-explanation*, *analogical reasoning*, and *other forms of reasoning* not falling into the first two categories), as well as for *valence* of affective states (i.e., negative vs. positive affect). The data analysis revealed that type of meta-cognitive event significantly affects pupillary response, with pupil size being statistically significantly larger for *self-explanation* events than for *other forms of reasoning*. Affective valence also had a significant effect on pupillary response, with pupil size being statistically significantly smaller during expressions of negative affect than during expressions of positive affect. The analysis in Muldner et al. (2009) does not provide concrete suggestions on how pupillary response can be used in real time for detection of positive versus negative affect or different types of meta-cognitive events. However, the fact that an effect of these states on pupillary response was found indicates that pupillary response should be further investigated as an additional source of information for student modeling.

## Recommendations and Future Research

---

In this chapter, we have discussed existing research relevant to understand the value of eye-tracking data in student modeling for ITS. This research indicates that the potential of eye-tracking data for student modeling is substantial, because there is evidence that these data can provide information on relevant learner states at the behavioral, cognitive, meta-cognitive, and affective level. In particular, work by Gluck et al (2000), Conati and Merten (2007), Kardan and Conati (2013), and Bondareva et al. (2013) show explicitly that a learner's eyes sometimes reveal more about cognitive and meta-cognitive processes than "overt actions" in a tutor interface. It follows that eye-tracking has the potential to enrich "standard" learner modeling techniques (i.e., those tapping only the regular interaction data).

Further research, however, is necessary to uncover the full extent of this potential. Eye-tracking data have so far been used to direct the adaptive behavior of an ITS by capturing only simple gaze patterns indicating attention or lack thereof (e.g., Silbert et al., 2000; Anderson, 2002; Wang et al., 2006; D'Mello et al., 2012). Student models that leverage gaze data to capture higher level student states such as learning (Kardan and Conati, 2012; Kardan and Conati, 2013, Bondareva et al., 2013), meta-cognition in terms of self-explanation (Conati and Merten, 2007), and affect in term of motivation (Qu and Johnson, 2005) have been developed, validated in terms of accuracy, but not integrated in an ITS. Although it is encouraging that positive results in terms of ITS pedagogical effectiveness have already been obtained by relying on simple gaze patterns (D'Mello et al., 2012), the next step for research in this area will be to see

if and how ITS effectiveness can be improved by relying on more sophisticated gaze-enhanced student models.

Another relevant next step is to exploit some of the insights generated by research on offline analysis of gaze data described in this chapter, to extend the usage of gaze data in student modeling and ITSs. For instance, although the findings of Gluck et al. (2000) on lack of attention to an ITS's interventions were exploited in Anderson (2002) to devise an ITS that can track this lack of attention and react to it, the work of Muir and Conati (2012) on factors that affect attention to hints can be leveraged to further improve how an ITS can increase this attention. For instance, Muir and Conati found that *attitude toward receiving help* generates consistent patterns of attention to hints throughout the interaction with the Prime Climb edu-game (low attention for those who do not want help, higher attention for those who do). Thus, if a student model can “see” that the student is not attending to a hint and knows that the student has a negative attitude towards receiving help, it can employ strategies specifically designed to increase attention to hints in someone who does not like receiving help, as opposed to using generic prompts as in Anderson (2002). It would also be interesting to investigate if and how the results uncovered by Muir and Conati (2012) generalize to other types to edu-games and to ITSs at large, and whether other factors may affect attention to hints (e.g., affective state, cognitive overload). A similar analysis could also be done for gaining more detailed insights on which factors affect attention to OLMs in general, and to specific ways to visualize them, especially considering recent results on the impact of individual differences (e.g., perceptual abilities and visualization expertise) and on visualization effectiveness (e.g., Conati and Maclaren, 2009; Toker et al., 2012). Finally, the results in Muldner et al., (2009) indicate that further research should be devoted to investigating how to use information on pupil dilation in student modeling.

Given that research on eye-tracking and student modeling is at a very early stage, as demonstrated by the relatively short list of references at the end of this chapter (we included all relevant articles we could find), should authoring tools for ITSs, such as ASPIRE, CTAT, and GIFT, support the use of eye-tracking data? If so, how? The answer to these questions depends on whether one views the primary purpose of such tools to support ITS research or support development of deployment-ready systems. Both are legitimate purposes and truly versatile authoring tools would cover both. Given that resources for development are always limited, however, existing tools tend to be more oriented towards either one purpose or the other.

When supporting ITS research is a priority, supporting the use of eye-tracking data would be an interesting forward-looking feature for an ITS authoring tool such as GIFT. Given that no best practices for employing eye-tracking data in student modeling have emerged yet, the authoring tool should support rapid prototyping of different ways of building student models that leverage gaze data. This capability would be of tremendous help in studying how eye-tracking might enhance student modeling. A useful first step is to enable researchers to do offline analysis of eye-tracking data combined with other key data sources, such as tutor log data. At minimum, this would require syncing the different data streams so they share common time stamps. A good next step would be to create a versatile architecture that enables the student modeling module (and perhaps other key modules of the ITS) to have access, at run time, to data from an eye-tracker. Steichen et al. (2013) have recently completed an eye-gaze service architecture to address exactly this need for data at run time. Their system, called Eye Movement Data Analysis Toolkit in Real Time (EMDAT-RT), is a standalone application that can provide real-time eye-gaze statistics to third-party applications through a lightweight web service interface. A client application (e.g., an ITS) can simply place a request for eye-gaze analysis (either at regular intervals or specific times), to which the service responds with real-time statistics (calculated either starting from a specific start time or for a specific time window, e.g., the last 10 seconds). Their system integrates a feature-rich open-source eye-gaze analysis module (called EMDAT), capable of calculating numerous summative gaze statistics beyond those usually provided by the analysis packages that come with commercial eye-trackers. The application has been designed to be application-independent, and may therefore be reused for different



application domains and purposes, including ITSs. The system (EMDAT-RT) and the internal analysis module (EMDAT) are currently compatible with Tobii eye trackers and will be released as open-source packages soon. Since both offline and online processing require interfacing with an eye-tracker's low-level API, an important goal would also be to make these tools independent of specific eye-tracker models or manufacturers, to increase versatility.

## References

---

- Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, 26(1), 85-112. doi:10.1207/s15516709cog2601\_
- Azevedo, R., Behnagh, R., Duffy, M., Harley, J., Trevors, G.: Metacognition and self-regulated learning in student-centered learning environments. *Theoretical foundations of student-centered learning environments* (2nd ed.). 171–197 (2012).
- Bondareva, D., Conati, C., Feyzi-Behnagh, R., Harley, J., Azevedo R., Bouchet, F. (2013). Inferring Learning from Gaze Data during Interaction with an Environment to Support Self-Regulated Learning. In *Proceedings of AIED 2013, 16th International Conference on AI in Education*, Springer.
- Bull, S., Cooke, N. & Mabbott, A. (2007) Visual attention in open learner model presentations: an eye-tracking investigation. *Proceedings of UM 2007, Eleventh International Conference on User Modeling*, (pp. 187-196.), Springer.
- Conati C. & Maclaren H. (2008). Exploring the Role of Individual Differences in Information Visualization. *Proceeding of AVI 2008, International Working Conference on Advanced Visual Interfaces*, ACM Press.
- Conati C. & Merten C. (2007). Eye-Tracking for User Modeling in Exploratory Learning Environments: an Empirical Evaluation. *Knowledge Based Systems* , 20(6), Elsevier Science Publishers.
- D’Mello, S., Olney, A., Williams, C. & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal Human-Computer Studies*, 70(5), 370-389. Elsevier.
- Gluck, K. A., Anderson, J. R. & Douglass, S. (2000). Broader bandwidth in student modeling: What if ITS were “eye”TS? In *Proceedings of the 5th international conference on intelligent tutoring systems (ITS 2000)* (pp. 504-513). Springer-Verlag
- Johnson, W.L., Rickel, J.W. & Lester, J.C. (2000) Animated pedagogical agents: face-to-face interaction in interactive learning environments. *Int. Journal on Artificial Intelligence in Education*, 11, 47-78.
- Kardan, S. & Conati, C. (2012). Exploring Gaze Data for Determining User Learning with an Interactive Simulation. *Proceedings of UMAP 2012, the 20th International Conference on User Modeling, Adaptation, and Personalization*, (pp. 126-138). Springer LNCS 7379.
- Kardan, S. & Conati, C. (2013). Comparing and Combining Eye Gaze and Interface Actions for Determining User Learning with an Interactive Simulation. *Proceedings of UMAP 2013, the 21th International Conference on User Modeling, Adaptation, and Personalization*, Springer.
- Mathews, M., Mitrovic, A., Lin, B., Holland, J. & Churcher, N. (2012) Do your eyes give it away? Using eye-tracking data to understand students’ attitudes towards open student model representations. In *Proceedings of ITS 2012, 11th International Conference on Intelligent Tutoring Systems*, (pp. 424-429). Springer-Verlag, LCNS 7315.
- Muir, M. & Conati, C. (2012). An Analysis of Attention to Student-Adaptive Hints in an Educational Game. *Proceedings of ITS 2012, 11th International Conference of Intelligent Tutoring Systems* (pp. 112-122) Springer LNCS 7315.
- Muldner, K., Christopherson, R., Atkinson, R., Bursleson, W. (2009) Investigating the Utility of Eye-Tracking Information on Affect and Reasoning for User Modeling. *Proceedings of UMAP 2009, 17th International Conference on User Modeling, Adaptation, and Personalization*, (pp. 138–149). Springer-Verlag.

## Design Recommendations for Intelligent Tutoring Systems - Volume 1: Learner Modeling

- Qu, L. & Johnson, L. W. (2005). Detecting the Learner's Motivational States in An Interactive Learning Environment. In Proceedings of AIED 2005, International Conference on AI in Education, (pp. 547 – 554), IOS Press.
- Sibert, L., Gokturk, M. & Lavine, R. (2000): The reading assistant: eye gaze triggered auditory prompting for reading remediation. Proceedings of UIST 2000, 13th Annual ACM Symposium on User Interface Software and Technology (pp. 101-107). ACM Press.
- Steichen, B., Schmid, O., Conati, C. & Carenini, G. Seeing how you're looking - Using Real-Time Eye Gaze Data for User-Adaptive Visualization. Submitted to 1st International Workshop on User-Adaptive Visualization (WUAV 2013), in conjunction with the 21st Conference on User Modeling, Adaptation and Personalization (UMAP 2013).
- Toker, D., Conati, C., Carenini, G. & Haraty, M. (2012). Towards Adaptive Information Visualization: On the Influence of User Characteristics. Proceedings of UMAP 2012, the 20th International Conference on User Modeling, Adaptation, and Personalization (p 274-285). Springer LNCS 7379.
- Wang, H., Chignell, M. & Ishizuka, M. (2006) Empathic tutoring software agents using real-time eye-tracking. Proceeding of ETRA 2006, Symposium on eye-tracking research and applications (pp. 73-78). ACM Press.