Abaris: Evaluating Automated Capture Applied to Structured Autism Interventions

Julie A. Kientz,[1] Sebastian Boring,[1,2] Gregory D. Abowd,[1] and Gillian R. Hayes[1]

[1] College of Computing & GVU Center, Georgia Institute of Technology
Atlanta, GA, USA
{julie, abowd, gillian}@cc.gatech.edu

[2] University of Munich
Munich, Germany
boring@cip.ifi.lmu.de

Abstract. We present an example of an automated capture application which provides access to details of discrete trial training, a highly structured intervention therapy often used with developmentally disabled children. This domain presents an interesting case study for capture technology, because of the well-defined practices and the tradition of manual recording and review of materials. There is a strong motivation for therapists to review the rich record of therapy sessions that is made possible by recorded video, but acceptance hinges on minimal intrusion upon the human activities. To achieve that, we leverage several perception technologies that fit with the natural activities of the live experience and allow the creation of meaningful indices. We also critically explore the contribution various perception technologies have on the overall utility of the capture system.

1 Introduction

Despite the increasing popularity of capturing everyday life activities (e.g. [7], [11], [22], [20]), there are still very few published examinations of real use, fueling views that automated capture may not be a compelling capability. Part of the reason for this is the difficulty of finding a domain for which frequent access of captured activity is likely. By identifying a high-need access situation and creating a reliable capture system, exploration of interesting research questions is possible. In this paper, we present an example of a domain with one such high-need access requirement—evidence-based behavioral and academic interventions for developmentally disabled children. We will address how automated capture and access impacts users in this domain, specifically, a team of collaborating therapists. We also determine whether it is worthwhile to employ various perception technologies to understand the captured activity, or whether simpler heuristics for indexing into captured media suffice.

An evidence-based approach to intervention therapy attempts to use empirical data of past performance to inform future decisions. For example, in medicine it is widely acknowledged that such empirical evidence is important for determining progress and

M. Beigl et al. (Eds.): UbiComp 2005, LNCS 3660, pp. 323-339, 2005.

guiding treatment decisions, particularly evidence that covers a patient's health outside the doctor's or professional's office. However, there are many situations in education, medicine, and other fields where gathering such evidence is cumbersome, if not impossible. In previous work, researchers explored the potential of automated capture for the specific case of treating children with autism [13]. One example suggested was the support for a team of therapists conducting discrete trial training (DTT), an application of Applied Behavior Analysis methods. An initial capture and access prototype was developed as a technology probe to determine whether practitioners of DTT saw any promise. Based on that feedback, and our own extensive experience over the past year with DTT, we have developed Abaris, a complete capture and access system to support home-based DTT therapy. Our intent is to evaluate the impact of some specific technology decisions on this popular intervention therapy.

This paper presents the design, initial deployment and results, and evaluation of the Abaris system. We pay particular attention to the use of two specific perception technologies during capture, Anoto's digital pen and paper technology [1] and Nexidia's phonetic-based speech detection [3]. These technologies allowed us to create a capture system that was a minimal departure from the existing DTT practice of our team of therapists, increasing the chances for adoption. The choice of each technology is strongly motivated by an understanding of the structure and practice of DTT. Preliminary results show that Abaris has been very well received by our users, and we will discuss why we think it has been successful. More importantly, we will explain whether the perception technologies played a necessary role in that success. The digital pen and paper were critical to the success of the system. While the speech recognition system did provide a useful indexing service to individual segment of therapy sessions, we argue that simpler heuristics would likely have fared just as well.

The structure of the rest of this paper follows. After reviewing relevant related work in automated capture and technology support for DTT, we provide an overview of the specifics of DTT and the structure and communication needs of a therapy team for one child. We then summarize the goals of the initial study to evaluate an automated capture solution to support DTT. We describe Abaris and present the results of a four-month pilot deployment, focusing on the use of Abaris during routine collaborative meetings of the therapists. We then analyze the results of the deployment, with a particular focus on reasons for the success of Abaris and an evaluation of the usefulness of the perception technologies as applied to this domain.

2 Related Work

In Ubiquitous Computing, Abaris falls into the category of automated capture and access applications. Several automated capture and access systems from research have helped explore this area, including applications for the classroom such as eClass/Classroom 2000 [7] and for meeting spaces such as Teamspace [22] and Tivoli [20]. Abaris provides users the ability to access information they already access manually (*e.g.,* notes in the classroom and meeting settings) supplemented with additional information to which they would not normally have access. These applications also focus on low need access situations, whereas Abaris is high need. Like Abaris, NoteLook [9], NotePals [10], and StuPad [24] all allow asynchronous annotation of videos in a collaborative setting, but are not designed for accessing multiple experiences as a collaborative activity. Abaris also differs from other capture and access systems, such as MyLifeBits [11], the Personal Audio Loop [14], Audio Notebook [23], and MIT's personal memory aid [25] in that these are designed for personal use of unstructured live experiences, rather than group access to a structured activity.

Other examples of technology that support the care of developmentally disabled children are not necessarily automated capture and access technologies. The child interacts directly with many, such as Simone Says for teaching speaking language skills [16] and the Discrete Trial Trainer [2], a commercial product aimed at simulated DTT. A commercial product known as mTrials [19], a PDA based data collection system for DTT, has enjoyed some success but lacks the additional audio and video inputs and collaborative access interface that Abaris provides.

We also relate our work to other ubiquitous computing projects that have focused on the design, authentic use, and/or summative evaluation of full-scale systems. PlantCareLabscape [5], mixed reality games [6], tour guides [8], Tivoli, and the eClass system are all end-to-end solutions to specific domain problems and were deployed for authentic use. Summative evaluation studies were published for all but the first. With the design and evaluation of Abaris, we demonstrate how acceptance and usability are impacted by specific perception technologies, similar to Benford *et al's.* assessment of the value of self-reported positioning in [6].

3 The Domain of Discrete Trial Training

The case study for this paper is based on a popular intervention therapy for children with autism, known as Discrete Trial Training (DTT). Developed in the 1970's by O. Ivar Lovaas [17], DTT has evolved as a specific method from the field of Applied Behavior Analysis (ABA) [4]. Though somewhat different from Lovaas' original conception, DTT is currently a best-practice method for teaching basic skills to children with autism and other developmental disabilities [15]. In DTT therapy, teams of trained therapists do one-on-one sessions with a child to teach basic skills in a structured setting. We outline the basics of DTT and the therapy team below. Although there is variation between different DTT practices, the description below is representative of standard practice, as implemented in a home setting.

The Discrete Trial. Advocates of DTT believe that even children with severe developmental disabilities can learn correct behaviors through controlled and conditioned training. A discrete trial is an example of this learning model. Once the therapist gains the attention of the child, she makes a direct verbal request to the child that requires a well-defined and correct response. If the child responds correctly, he is immediately rewarded with a reinforcing stimulus, such as a piece of candy, a favorite toy, or verbal praise. If the child responds incorrectly, the therapist prompts the child in a way to ensure a correct response. The trial is immediately repeated, with the therapist providing whatever prompt is needed to guarantee a correct response. The therapist records the result of the trial (*I* for an independent or correct response; otherwise, any of seven or eight letters that represent the prompting used by the therapist). If a “correction” trial follows the initial prompted response, the therapist may also record the result of that correction trial.

A DTT Program. The therapy regime for DTT consists of a collection (10 to 20) of programs for which data is collected. Each program consists of a basic skill (*e.g.*, Picture Identification), a target (*e.g.*, picture of a dog), a note further explaining the task (*e.g.*, selection from a field of three pictures), and a specific command (*e.g.*, “Give me the <target_name>.”).

A Therapy Session. Each program/target combination is performed a number of times, ideally distributed randomly throughout the 1-2 hour session. All data is recorded on a scoresheet. A graph, as shown in Figure 1, indicating progress for each program is updated at the end of the session to reflect that day’s data.

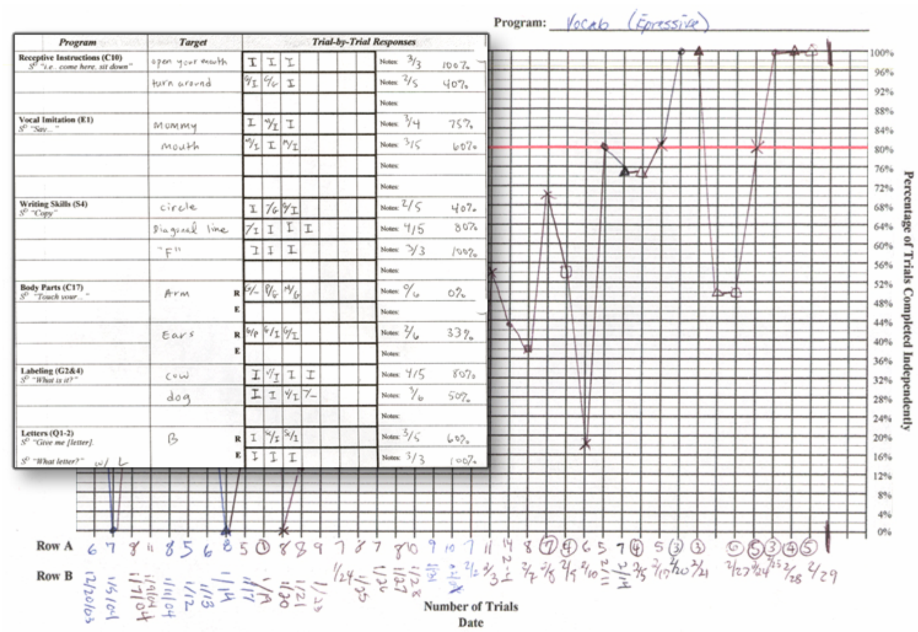

Figure 1: Examples of paper-based forms used by therapists. Left is a therapy data sheet completed during therapy, and right is a graph of the child’s progress for a particular skill.

Advancing the Program. A given program/target combination is “mastered” when some pre-defined performance level (*e.g.*, 80% correct responses on a given day) is achieved over some interval of time (*e.g.*, three consecutive days). Once a program/target combination is mastered, the target is changed. When a sufficient variety is mastered for a program, the program is mastered overall. Mastered program/target combinations are practiced (without data collection) throughout a therapy session.

Before a session, a therapist reviews the child’s therapy materials. She consults a notebook containing the child’s past session data sheets, program progress graphs, mastered skills, and narrative notes from other therapists on the child’s progress. She reads over the notes written by other therapists and prepares her session materials, which includes pictures, objects, and writing utensils. After she has prepared everything, she begins the session with the child by playing and interacting with him, and then brings him to the table to rehearse mastered skills and work on target skills.

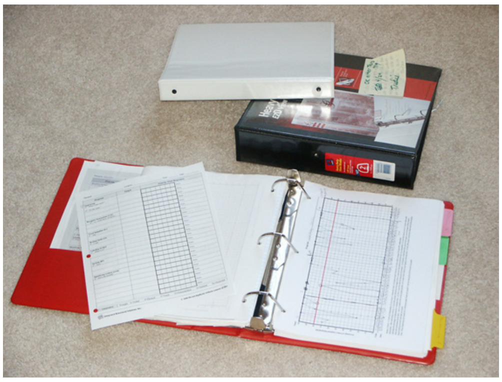

Figure 2: The left image shows a therapist engaged in a therapy session with a child. The right image shows an example of the large amounts of paper required to store all of the data.

Evaluating Progress. The team we studied consisted of a parent (trained in DTT but not practicing), three regular therapists, one lead therapist, and a consultant, all providing therapy to a seven year-old, low-functioning child diagnosed with Autistic Disorder (mild to moderate), using the DTT procedure described above.[1] The lead therapist has additional tasks of administrative paperwork, such as determination of which program/target combinations are mastered and scheduling new targets and programs for future sessions. The consultant does no direct therapy with the child, but is an expert in behavior analysis. The team typically meets every other week to discuss therapy, analyze data, and make any necessary adjustments to help the child learn the skills more effectively. The consultant leads these meetings and uses the manually recorded data as an agenda to run the meeting (see Figure 1). The consultant looks at the book of graphs and asks the therapists for details on how the child is progressing on each skill. If a certain target skill has been in place for a long time with little improvement, the team may remove the target and replace it with another one, or they may discuss why they do not think he is learning it. Therapists try to remember details of what occurred in their sessions and make hypotheses about what is causing him to perform particularly well or poorly. The consultant will make suggestions with the team generally implementing these within the next two weeks of therapy. After making these changes, the team reviews the progress again at the next meeting.

[1] Two members of the team were from the research team reporting this work. All other team members and the child are protected as human subjects under an approved IRB protocol.

4 Research Goals

With a firm understanding of the domain, we now frame the goals of our research into the design and evaluation of Abaris. We begin with two observations:

- The therapy sessions, though fairly fast-paced and flexible, have a well-defined structure that can be leveraged naturally by perception technology, potentially providing a suitably indexed recording for later access.
- The team meetings present a high-need example of access, in which the users who are both capturing and accessing the data absolutely require it to perform their jobs. Furthermore, the meetings consist of a lot of self-reported reflections on past experience between therapist and child, a clear opportunity for improvement with real evidence of what transpired during a therapy session.

Although DTT therapy is a relatively well structured and successful treatment for children with autism, there are some deficiencies in the process that may lead to inaccuracies in the interpretation of the data, making the overall therapy less efficient. Our goals for the Abaris system were to be able to address some of these issues and make therapy a better and more useful experience for the therapists and the child.

Discrete trial training is particularly well suited to the use of automatic capture technologies. Therapists and parents alike are highly motivated to use anything that will save time on laborious paperwork yet does not reduce the quality of the intervention. Additionally, it is a structured activity, with individuals already trained to be cooperative in the process of manually recording data. Because therapists record, calculate, and graph all of the data by hand, there is a high likelihood that the data may be inaccurate due to simple human error. Furthermore, graphing and calculating all of the data using pen and paper is time consuming, often requiring up to one third of the session and taking time away from the child's instruction. By designing a system that automates a lot of this hand analysis and calculation, we can reduce the amount of time spent in paperwork, similar to how others have found that automation can save time in paperwork for Activities for Daily Living (ADLs) [21].

DTT requires a significant amount of improvisation and thus we must design Abaris to be as flexible as possible when capturing data from the therapists. Since pen and paper allows anything to be written anywhere on the page, we feel that keeping the paper for capture is essential. Besides flexibility, we believe that allowing therapists to keep a pen and paper system will have minimal change to existing practice, which will increase the likelihood of acceptance, as noted by Mackay *et al.* in their study of air traffic controllers [18]. The challenge is to design capture in a way that maximizes the inherent structure of sessions without violating the process, and to provide a nimble access interface that would encourage exploration of the evidence without requiring too much time, effort, and distraction during team meetings.

At current team meetings, therapists speculate about whether a child is responding to prompts in certain ways, how well the child is focused, whether or not the child exhibits some affect, and whether the therapist is conducting each trial correctly. Much of the grading of each trial is subjective, especially in the grading of word pronunciation or letter formation, thus discrepancies in the grading of the child by multiple therapists tend to interfere with measures of progress. These discrepancies can lead to a mismatch in skills taught and the child's abilities, which can be frustrating both for the child and for the therapist. Capture of rich data, such as video, allows therapists to *see* what each of the other therapists is discussing without being present during therapy sessions and to notice things that the therapist herself may not have, ensuring increased consistency and enabling more accurate decisions and advice. Thus, the overall goal for the access interfaces is to provide a means of facilitating discussion amongst groups of therapists about trends in the data using easy access to both empirical and rich data to enable data-based decisions for long-term use.

5 The Abaris System

Abaris contains two major software components—one for *capture* or recording of data, and one for *access* and analysis of data—which are located on the same computer. This computer can act as a network server to allow remote use for certain tasks, like maintaining the programs and viewing the captured sessions. As shown in Figure 3, additional devices supplement the software on the single PC including a high quality printer for augmented datasheets, a web cam for capturing video and audio data, a high-quality wireless microphone for voice recognition, and a digital pen for writing the grades on the specially printed paper.

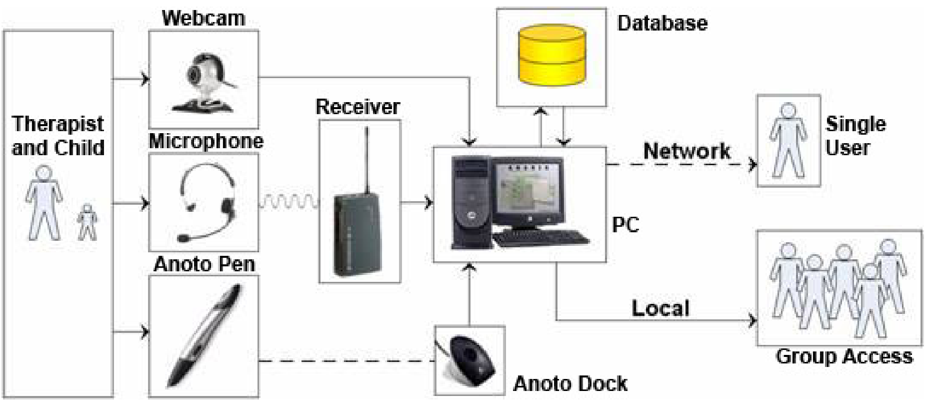

Figure 3: This shows the basic system setup to run and interact with Abaris.

5.1 The Capture Interface

Recorded video from therapy sessions coupled with appropriate indexing allows fast access to particular trials. In current practice, therapists use both a spoken command to indicate the beginning of the trial to the child and a pen to record data after a trial. We leverage these practices to create effective indices into the captured therapy session. Using Nexidia's voice recognition technology (off-the-shelf, phoneme-based speech system), we can retrieve timestamps for a specific command, obtaining estimates for trial beginnings. After trials, therapists record grades on the augmented datasheet using a special pen (see Figure 4, center). Replacing traditional pen and paper with Anoto's digital pen technology affords collection of positions and timestamps of every stroke, while preserving the flexibility inherent to writing.

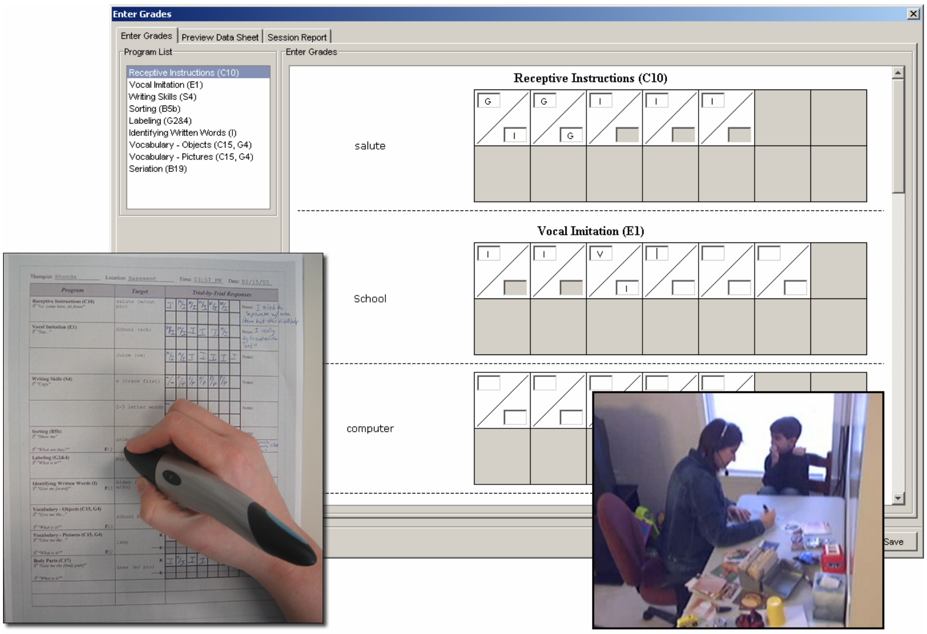

Figure 4: Left is the digital pen and its specially designed paper used for entering trial grades. Center shows the current system does not use any character recognition, instead providing an interface to enter grades. Right is a scene from a therapy session using the Abaris system.

While capturing a session, Abaris records an additional audio file, which is monitored and indexed by Nexidia while recording, including a pattern file that can be searched for speech patterns indicating the beginning of a trial. Within the plain-text XML file generated by the digital pen's interaction with a data sheet, each stroke is stored with its coordinates and associated absolute beginning timestamp. A stroke, by definition, contains at least 6 pixels and more than half of its points inside the 31x20-pixel cell the system is analyzing, preventing erroneous marks on the paper made by therapists signaling trial data. Using data stored from the written records and the patterns in the audio, Abaris reconstructs likely beginning and ending times for particular trials. Further analysis on these timestamps is discussed in the following sections.

5.2 Access Interface

The access interface for Abaris provides therapists with the ability to review sessions as well as to correct grades and timestamps for places where technical or human error created incorrect data. Therapists need to perform these tasks both locally at the site of therapy and remotely from their homes or offices in preparation for team meetings and therapy sessions. Furthermore, they must be able to access Abaris both individually and in a group setting during team meetings. Of course, Abaris must provide at least the same level of functionality as the traditional pen and paper process, including graphing of empirical data and review interfaces for therapist datasheets.

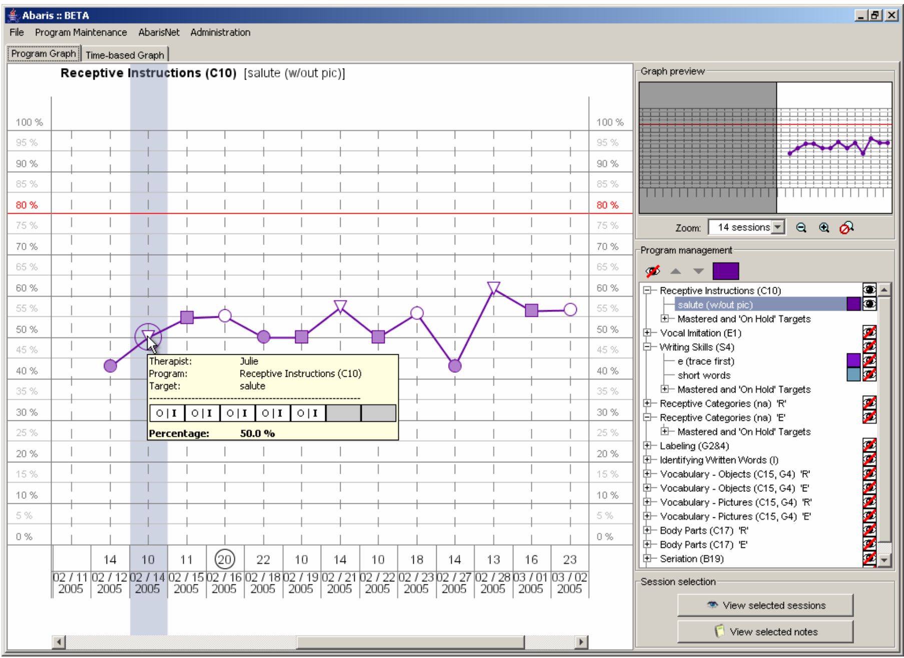

Figure 5: This shows the main access interface displaying a single selected graph on the left, with tool tip indicating information for a specific session. The right shows a view of the entire graph and the list of selectable programs.

Once the access interface is started, the therapist/consultant can choose which programs to view by marking programs and targets to be shown or not. If more than one target is visible the graphs are overlaid in the same view with a displayed legend. Because multiple graphs might become confusing, other visualization techniques facilitate analysis. A tool tip (describing the target and program) appears each time the cursor is near a target's line. Another tool tip shows the data of a target from a particular session when the user hovers near that data point. Figure 5 shows an example of a typical graph with a therapist specific tool tip.

Users can select multiple sessions for which they want to view more details by clicking and highlighting the columns associated with those sessions. This functionality allows the user to review two different sessions quickly to compare procedures. The session browser loads in its own window, with typical video control functions of play, pause, stop, fast forward, and frame seeking functions as well as functions to jump to the next or previous trial of currently visible programs. Along the bottom of the window is a zoomable timeline that shows when trials occurred, using the predictions described above. To the right of the video are the grades for selected programs. Clicking on a grade moves the video to the start time for that trial. If there are several sessions loaded, a user can switch between them by clicking on the timeline of another video or selecting a trial that is not part of the video currently shown. The grades assigned to a trial, as well as the beginning and end times, can be modified. These corrections appear on the graphs immediately after saving the changes. Within the access interface, therapists can also add and edit programs and targets, an activity that happens frequently during the course of a team meeting.

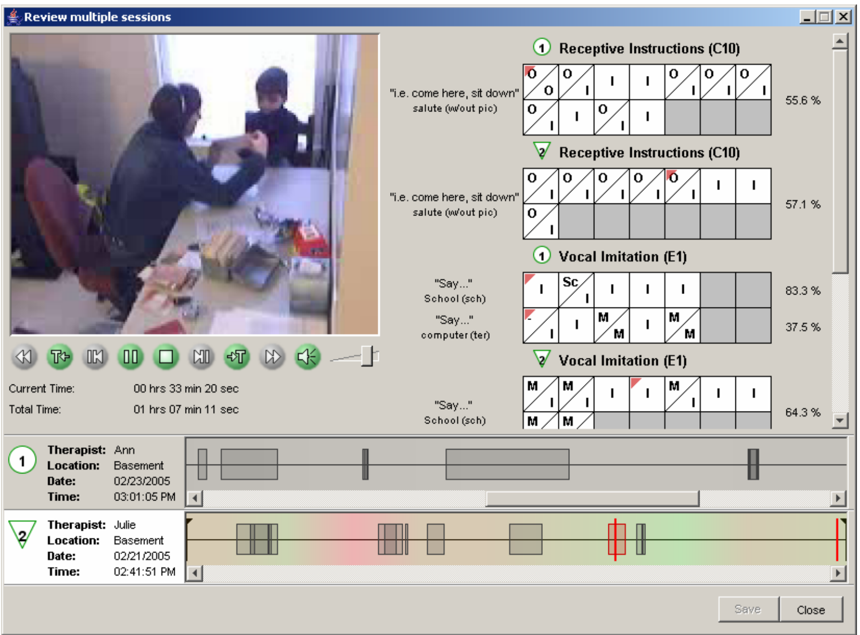

Figure 6: This shows a session browser set up to view two different therapy sessions. The bottom shows the two different timelines, and the left shows the grades for the different trials.

6 Results of Deployment

We deployed the Abaris system described above in one child's home for his team of therapists to use in a pilot study for four months. Meetings to discuss the progress of the child approximately every two weeks for a total of six meetings. We instructed therapists on use of the capture system before deployment, and a researcher was on hand during their first sessions to answer questions. Before the first meeting, we trained the lead therapist and the consultant to use the access interface. During the meetings, researchers were present to answer questions. In the first meeting, one of the researchers controlled the access interface according to therapist requests; at subsequent meetings, the lead therapist controlled the interface.

Our main goals for Abaris were to reduce the amount of paperwork time during sessions, obtain accurate timestamps to allow therapists to index into videos of certain trials, evaluate the use of the perception technologies, determine the usefulness and usability of the Abaris system, and facilitate group meeting discussions.

6.1 Use of Capture System

To date, the team has captured 52 sessions, consisting of 3869 trials and 45.1 hours of recorded data, including every session that has taken place during the study. The capture interface was easy for the therapists to learn, because the digital pen allowed therapists to perform their work in the exact same way they had done it before. Although the interface appeared to be easy for therapists to use, they initially demonstrated skeptical attitudes about its use. Despite this skepticism, participants used Abaris in all of their sessions for which it was available. The only benefit of use at this stage was removing the need for users to "hand graph." This consistent use is remarkable given that at first, all users were contributing to this groupware system while receiving little benefit [12]. We believe this was due in large part to the conscious effort during design to maintain nearly identical work practices that reduce or maintain the same level of effort. At the first meeting that made use of Abaris, participants were then able to experience the benefits of access.

Therapists reported allocation of session time both before and during the deployment. Overall, work time for these hourly employees decreased slightly, but this may be the result of fewer target skills for the child during the time of the deployment due to the child being sick or having a difficult time in school. The percentage of time that therapists spent in paperwork decreased, resulting in more time spent teaching or playing with the child (see Figure 7). Thus, with Abaris, therapists can devote a greater percentage of their paid time to interaction with the child.

Two therapists reported that the clip-on microphone was a bit too heavy for some of their typical clothing and could be uncomfortable. Most preferred to use a head-mounted boom microphone. A few incidents occurred in which the child became fascinated by the microphone and would reach out and play with it, a behavior that typically occurs when therapists wear jewelry the child finds interesting. Although this behavior can be common for some children with autism, it may not happen in all cases. We considered using a room level microphone, but the child often vocalizes during therapy sessions, which affects the accuracy of the voice recognition.

Simple usage errors sometimes had large impact. One of the therapists forgot to press the record button at the beginning of her session, resulting in no video for the session. In one incident, placement of the Anoto paper in the printer backwards resulted in incorrect detection of the timestamps. These errors can be prevented future versions of Abaris, but because of its improvisational nature, we could not predict all of the exceptions to the therapy. For example, the lead therapist wanted to change the success criteria for one type of program, but she had no way of doing this with our current interface. Basing Abaris on pen and paper input allowed for a lot of improvisation, but it was very difficult to plan and address all cases.

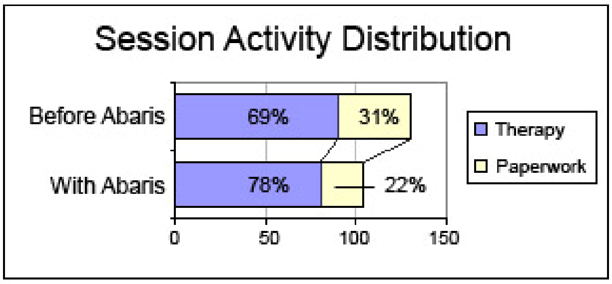

Figure 7: Comparison of the activity makeup of each session before and after deployment.

6.2 Use of Access System

Therapists used the access interface for discussion during two meetings, which lasted 2.5 hours and 1.5 hours, respectively. Each meeting was video recorded and observed, and afterwards we debriefed the therapists on the experience with the system, in which discussion was similar to that of a focus group. Between the first and second meetings, we instrumented the access interface so that we could produce logs of its use providing some empirical evidence of access behaviors. In the second meeting, the team used the access interface to view the video six times, and video viewing took up 20.4 percent of the meeting time. Visualizations of interesting data in these logs are present in Figure 8. The top graph is a typical example of comparing a program across two therapists viewed by the lead therapist before the meeting. The middle graph shows various artifacts in the interface—the timeline and the trial grades—were used to navigate to the desired portion of video. The bottom graph is a detailed version of a portion of the middle graph. That this kind of browsing occurred six times during the meeting is an indication that the team found the value of viewing video outweighed the cost of finding the appropriate session. For 18 months prior to these two meetings with Abaris, the team had access to digital video recordings of the sessions at the site of the meeting, and not once was a segment viewed during a meeting, reportedly because it took too long to find a relevant clip.

Due to the complexity of the data recorded for DTT, therapists reported the access interface to be complicated at first. They received two hours of training before expressing enough comfort to use it on their own. Although ease of use was not as good as we would have liked, therapists reported that the benefits of the system were worth the time it took to learn the access interface. Additionally, the access interface is intended for expert users (*e.g.* the lead therapist and the consultant), allowing them to use the system with all of their clients once they are past this initial learning curve.

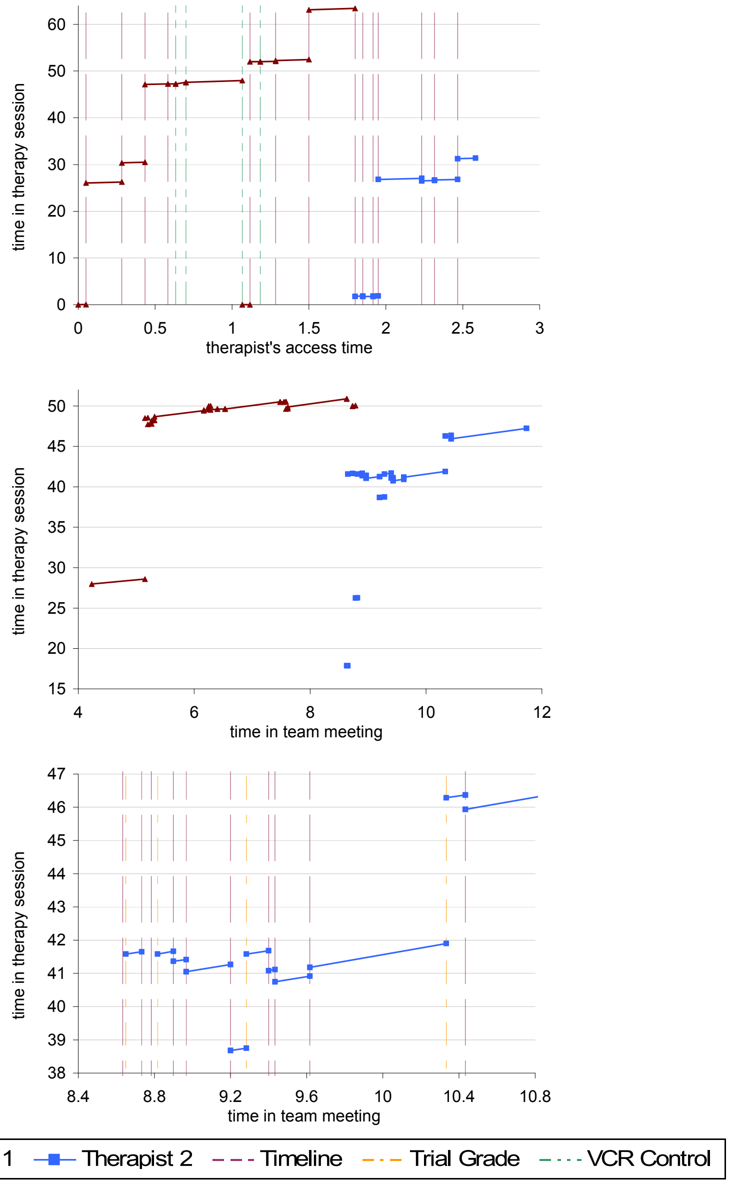

Figure 8: Visualizations from the logs of the second meeting of the team. Top shows access from the lead therapist before the meeting while the center graph shows access to videos during the meeting. The plot with dark, triangle data points shows access to video of one therapist, while the plot with light, square data points shows viewing a different therapist. In the bottom graph, we expand one segment of the top graph (between 8.4-10.4 minutes into the meeting) and show how different artifacts are used to facilitate navigation.

6.3 Overall Therapist Impressions

Though we are still in the process of fully understanding the impact that Abaris had on our team we deployed with, it should be noted that Abaris has been very well received by our team. Through initial post deployment interviews, we have discovered that the only complaints on the capture side were that the microphone was distracting to both the child and to the therapist. Therapists felt that the meetings were more structured and Abaris helped to create an agenda for the meeting. The therapists and the consultant both commented on how helpful it was to see how others did therapy so they could verify that they were doing things correctly. It also allowed them to all see the data, which helped in facilitating the discussion of progress. Aside from the occasional glitches in the code of Abaris, the therapists were very excited about deployment and all of them stated they were sorry that the deployment had to end. We are continuing to assess the impact of the system on the overall therapy sessions and team meetings by observing and analyzing how the team operates post deployment.

7 Discussion: Do Perception Technologies Make a Difference?

The fact that Abaris is considered a useful system by its target user group is encouraging, but as researchers and designers, we want to better understand which features contribute to its usefulness and which do not. The integration of trial time predictions and the recorded video are a reasonable first guess at the success of Abaris. As seen in Figure 8, skimming to an appropriate portion of the video was quick enough to encourage use. End times of trials were equated with the time the grade for that trial was written on the Anoto paper. Beginning trials were estimated based on suggested locations of the appropriate verbal command. We selected four separate therapy sessions, one for each therapist, and used Abaris to create "ground truth" timestamps for the beginning and ending of each trial by manually noting when trials began. Figure 9 shows the error distribution of prediction versus ground truth. A negative error indicates a time prediction earlier than ground truth, and a positive error indicates a prediction after ground truth.

For each of the programs, the error distribution of the Anoto predictions is much narrower than that for Nexidia. The Anoto predictions occurred temporally after the actual end time, as expected, because trials are graded after they occur. The distribution of errors for Nexidia is wider. When viewed grouped by therapist, these error distributions have substantial variation in practice between therapists. Therapst 2's Anoto predictions were very tightly bunched near the actual end of trials. This therapist followed the practice of writing trial grades right after the trial was performed, as opposed to other therapists who ensured delivery of a reinforcing reward first. This is actually considered good practice for DTT, and Abaris benefits from this practice.

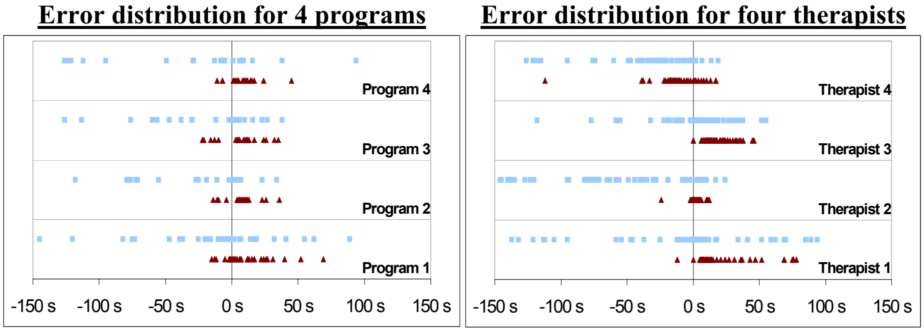

Figure 9: On the left, paired error distributions (in seconds) for Anoto-predicted end of trial (dark triangles) and Nexidia-predicted beginning of trial (light squares) for four of the programs used in the deployment. On the right, error distributions are shown for one session of each of the four therapists (lead and three regulars).

The phoneme-detection of Nexidia, and our accompanying algorithm for assigning assumed beginning of trial times, produced a significant amount of error. Errors are not surprising, given the nature of the therapy, with graded and mastered trials often having the same spoken command and occurring in rapid succession. However, because the interface was still usable, as reported based on use during team meetings and the overwhelming positive reaction of the team of therapists in discussions, this error may not be limiting. If this size of error makes no discernable difference, we hypothesize that speech detection may be unnecessary if we can find an alternative approach that introduces no additional errors.

Unfortunately, voice recognition only provides a best guess for the beginning of a particular trial, because many trials for which grades are not recorded use the same spoken command. For example, a therapist may be grading a child's ability to mimic hand clapping, for which the spoken command is "do this" coupled with the therapist modeling hand clapping. Prior to this trial of interest, a therapist may ask the child to perform any number of other activities with the same command of "do this", and then end with the final request "do this" while hand clapping. Thus, we considered a vision-based solution in which therapists used a simple two-finger gesture on a score sheet to indicate the beginning of a graded trial and the actual grades before and after the trial itself. Though the approach was simple to teach and the vision problem was feasible, we found that therapists could not remember to do the gestures at the correct times, resulting in a loss of grading information. Instead, we developed a simple algorithm for determining the most likely beginning of a trial based on a combination of the time that the trial likely ended (from the Anoto data) and the time that different spoken commands were used.

Considering the narrow distribution of the Anoto errors for trial endings in Figure 9, there are several suggestions for potential temporal heuristics that might produce begin trial estimates at least as good as Nexidia. We have anecdotal evidence that for Therapist 2, a fairly reliable heuristic was a function of the program type and whether or not a correction trial was needed. Our current results give us confidence that we know an upper bound on the error distribution for estimating the start of a trial, and we will experiment with a variety of algorithms to find one that is both accurate and precise enough without impinging on the therapy itself.

8 Conclusions and Future Work

We presented our work on the design, development, and deployment of the Abaris system for supporting therapists who do discrete trial training therapy for children with autism. Our results show that initial therapist reaction is largely positive, which we attribute to the closeness of the system to the therapists current practice. We improved the practice by allowing therapists to spend less time in paperwork and more time in the therapy itself. Also, we evaluated the usefulness of the recognition techniques employed by Abaris by comparing its accuracy with the ability to find the needed video segments. Time stamping using Anoto digital pen technology was useful for this practice, and while the errors were introduced by using Nexidia voice recognition, the indices to the video still were useful in practice. Thus, we hypothesize that you could do just as well as the voice recognition with a trained heuristic based on therapist and program type.

Though our initial results are promising and have lead to some interesting insight, there is still room for exploration with this technology. As of the time of writing, Abaris has just finished the four month deployment, and we are in the beginning stages of analyzing the massive amount of data we have collected beyond what is presented in this study. We plan to take the feedback we've received from the pilot and make final improvements to the system and deploy it with a new team of therapists of which the researchers are not members. Plans for the Abaris system include more visualization of data that otherwise wouldn't be possible with the paper system, using it as a test bed for more recognition, automation, and multimodal interaction techniques and finding ways of sharing a child's therapy information with all those interested in his progress, not just those present at therapy meetings, such as through a web based information portal. Lastly, we are hoping we can use Abaris to contribute to the field of autism interventions by enabling domain experts to analyze the science behind DTT therapy itself and improve on its methods.

Acknowledgments

This work was sponsored in part by the National Science Foundation (ITR grant 0121661), the Aware Home Research Initiative, and the Cure Autism Now foundation. The authors wish to thank Integrated Behavioral Solutions Consulting and our team of therapists for supporting and evaluating us in this work. We also wish to thank Albrecht Schmidt at the University of Munich, Juane Helfin at Georgia State University, Francois Guimbretière and Dave Levin at the University of Maryland, and Georgia Tech's Ubicomp, Autism, and IMTC research groups.

References

1. *Anoto, Inc. Website*, in *http://www.anoto.com*. 2005.
2. *Discrete Trial Trainer (DTT)*, in *http://www.dttrainer.com*. 2004: Columbia, SC.
3. *Nexidia, Inc. Website*, in *http://www.nexidia.com*. 2005.
4. Alberto, P.A. and A.C. Troutman, *Applied Behavior Analysis for Teachers*. 6th ed. 2003: Prentice Hall.
5. Arnstein, L., et al., *Labscape: A Smart Environment for the Cell Biology Laboratory*, in *IEEE Pervasive Computing Magazine*. 2002.
6. Benford, S., et al. *The error of our ways: The experience of self-reported positioning in a location-based game.* in *Ubicomp 2004*. 2004. Nottingham, UK: Springer-Verlag.
7. Brotherton, J.A. and G.D. Abowd, *Lessons learned from eClass: Assessing automated capture and access in the classroom.* ACM Transactions on Computer-Human Interaction, 2003.
8. Cheverst, K., et al. *Developing a context-aware electronic tourist guide: some issues and experiences*. in *SIGCHI conference on Human factors in computing systems*. 2000.
9. Chiu, P.I. *NoteLook: Taking Notes in Meetings with Digital Video and Ink*. in *ACM Multimedia*. 1999. Orlando, FL.
10. Davis, R.C., et al. *NotePals: Lightweight Note Sharing by the Group, for the Group.* in *CHI 1999*. 1999. Pittsburgh, PA.
11. Gemmell, J., et al. *MyLifeBits: Fulfilling the Memex Vision*. in *ACM Multimedia '02*. 2002.
12. Grudin, J., *Groupware and Social Dynamics: Eight Challenges for Developers.* Communications of the ACM, 1994.
13. Hayes, G.R., et al. *Designing Capture Applications to Support the Education of Children with Autism*. in *Ubicomp 2004*. 2004. Nottingham, UK.
14. Hayes, G.R., et al. *The Personal Audio Loop: Designing a Ubiquitous Audio-Based Memory Aid*. in *Mobile HCI 2004*. 2004. Glasgow, Scotland.
15. Heflin, L.J. and R.L. Simpson, *Interventions for Children and Youth with Autism: Prudent Choices in a World of Exaggerated Claims and Empty Promises. Part I: Intervention and Treatment Option Review.* Focus on Autism and Other Developmental Disabilities, 1998. **13**(4): p. 194-211.
16. Lehman, J.F. *Toward the use of speech and natural language technology in intervention for a language-disordered population*. in *Third International ACM Conference on Assistive Technologies*. 1998.
17. Lovaas, O.I., *Teaching Developmentally Disabled Children: The Me Book*. 1981, Austin, Texas: Pro-Ed.
18. Mackay, W.E., et al. *Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control*. in *CHI 1998*. 1998. Los Angeles, CA.
19. Mobile Thinking, I., *mTrials*. 2002: San Diego, CA.
20. Pedersen, E.R., et al. *Tivoli: An Electronic Whiteboard for Informal Workgroup Meetings*. in *ACM INTERCHI 1993*. 1993. Amsterdam, The Netherlands.
21. Philipose, M., et al., *Inferring activities from interactions with objects*, in *IEEE Pervasive Computing*. 2004. p. 50 - 57.
22. Richter, H., et al. *Integrating Meeting Capture within a Collaborative Team Environment*. in *ACM Conference on Ubiquitous Computing*. 2001. Atlanta, GA.
23. Stifelman, L.J., *The Audio Notebook*, in *Media Laboratory, MIT*. 1997.
24. Truong, K.N. and G.D. Abowd. *StuPad: integrating student notes with class lectures*. in *CHI 1999 Extended Abstracts*. 1999.
25. Vemuri, S., et al. *An Audio-Based Personal Memory Aid*. in *Ubicomp 2004*. 2004. Nottingham, UK.