KATHARINE BLOCHER, ROSALIND W. PICARD

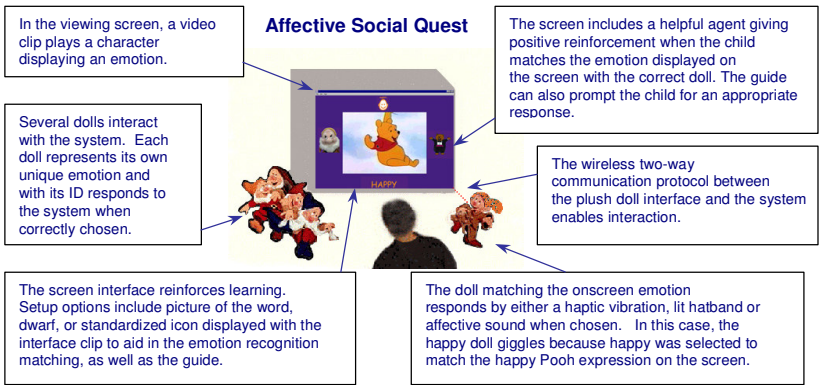
AFFECTIVE SOCIAL QUEST

Emotion Recognition Therapy for Autistic Children

1. INTRODUCTION

Recognising and expressing affect is a vital part of social participation. Unfortunately, those with autism have a learning disability in this area, often accompanied by deficits in language, motor and perceptual development. Their development of social communication is very low compared to neurologically typical children who learn social cues naturally while growing up. In trying to comprehend social nuances in communication or social behaviour to blend in during everyday interaction, autistic children get frustrated, not only with themselves but with their teachers too, and often give up learning. What may help an autistic child in this case is an ever-patient teacher.

This research presents an approach to creating that teacher: a persistent and unresentful aid that progressively introduces basic emotional expressions, guides recognition development through matching, and records the child's success. It is designed to teach emotion recognition to autistic children with a heterogeneous disorder. Although the application developed for this research does not come close to the abilities of a highly trained human practitioner, it is designed to offload some of the more tedious parts of the work.



Kirsten Dautenhahn (ed.), Affective Agents, 1—8.

Figure 1. Elements of the Interface

Affective Social Quest (ASQ) (Figure 1) consists of a computer, custom software, and toy-like objects through which the child communicates to the computer. The system synthesises interactive social situations in order to promote the recognition of affective information. This system will not tire because of impatience and can be a safe place for the child to explore. The goal of ASQ is to provide an engaging environment to help children -- specifically autistic children – learn to recognise social displays of affect.

Affective computing is an area of research aimed at giving computers skills of emotional intelligence, including the ability to recognise and respond intelligently to emotion (Picard 97). A computer can be taught to recognize certain aspects of emotion expression, such as facial movements indicative of a smile, and can prompt people for information related to human emotional state. However, computers are limited in their ability to recognize naturally occurring emotions; they can not easily generalize patterns from one situation to the next, nor do they understand the emotional significance associated with emotion expression. We recognize that some of the main problems we face in trying to give computers emotion recognition abilities are similar to those faced by therapists trying to help autistic children learn to recognize emotion. We expect that progress in either of these two areas will help inform progress in the other.

Six discrete emotions that show up universally with characteristic facial expressions are: happiness, sadness, anger, fear, surprise, and disgust (Ekman 92). We chose four of these emotions for ASQ: happiness (happy), sadness (sad), anger, and surprise. ASQ can display the emotion word, icon, doll face and representative video clips for the autistic child. The aim is to help the child scaffold different representations for an emotion, to help him or her generalise the many ways one emotion may be displayed.

Different approaches for behaviour intervention are available to parents of autistic children. Many programs instruct children using emotion words and icon representations, showing children photographs of people exhibiting emotional expressions using learning development aid emotion cards. However, systematic observations or experimental investigations of specific social behaviours are few (Blocher 99, Lubin 99, Tuchman 99, Sigman & Capps 97). One example of emotion instruction is Howlin and Baron-Cohen's practical guide to teaching emotions and mental states to children, *Teaching Children with Autism to Mind-Read 1999*, which may soon be available on CD-ROM.

Many children with autism are drawn to computers, and can become engaged with off-the-shelf software. Software tools that augment the child's training are available. Most applications focus on verbal development, object matching, or event sequencing. Laurette software is designed for autistic children to solve 'what if' scenarios and help them decide what the next action in a sequence could be. Mayer-Johnson has a "board maker" software tool that combines words with its standardized icons (Picture Communication Symbols (PCS)), to help children communicate through pictures.

Why ASQ? *Affective Social Quest* builds on the strengths of autistic children's visual systems through use of video. Additionally, it incorporates characteristics of the intervention methods listed earlier. Different displays of emotional expressions presented in moving pictures may be helpful to autistic children for seeing the frame-by-frame development of an emotional expression. The potential for using affective computing and physical interfaces in therapy forms the heart of this work.

2. THE SYSTEM

ASQ displays an animated show and offers pedagogical picture cues -- the face of the plush dwarf doll, the emotion word, and the Mayer-Johnson standard icon -- as well as an online guide that provides audio prompts to encourage appropriate response behaviour from the child. The task was to have the system act as an ever-patient teacher. This led to a design focused on modeling antecedent interventions used in operant behaviour conditioning. In essence, ASQ represented an automated discrete trial intervention tool used in behaviour modification for teaching emotion recognition.

The system has multiple possibilities for interaction. In the default case, the system starts with a video clip displaying a scene with a primary emotion (antecedent) for the child to identify and match with the appropriate doll (target behaviour). After a short clip plays, ASQ returns to a location in the clip and freezes on the image frame that reinforces the emotion that the child is prompted to select. The child is then prompted to indicate which emotion s/he recognizes in the clip, or frame – i.e., to select the appropriate doll matching that expression. To motivate interaction, the doll interface -- the only input device to the system -- creates a playful interaction for the child.

The practitioner can, using various windows, customize each interaction for each child. First, the practitioner can choose which video clips the child will be shown. These clips are arranged based on content (e.g. Cinderella), source (e.g. Animation), complexity (low-med-high difficulty of recognizing the emotion), duration (clip length), and emotion (happy, angry, sad, surprised). Practitioners may wish to focus on only a couple emotions early on, or may wish to avoid certain types of content depending on the idiosyncrasies of a particular child. The child's interface screen can be configured to include one or all of the following picture aids: Mayer-Johnson standardised icons representing the emotion, the word for that emotion, and a picture of the doll's (emotional) face. Also, an optional online animated guide can be included on the screen; the guide can be configured to provide an audible prompt (discriminative stimuli) for child interaction, or verbal reinforcement (putative response) for child affirmation.

The practitioner can configure many kinds of cues for ASQ to use in aiding the child. The dolls can cue the child with one of the following three choices: affect sound, hatband and lights, or internal vibration. The system can be cued to audibly play one of three different sequences to prompt the child to select a doll to match the emotion in the video clip: for instance, when a *happy* clip plays, the system will say, "*MATCH HAPPY*" when ***Match*** is chosen, or say, "*PUT WITH SAME*" when ***Same*** is

chosen, or say, "*TOUCH HAPPY*" when ***Touch*** is chosen. Likewise, reinforcements for incorrect doll selections will say, "*THAT'S SAD, MATCH HAPPY*" for ***Match***. Seven different cue set-ups are configurable for one session with the timing, sequence, and repeat rate tailored for each. Thus, the practitioner can control exactly what is said to the child. The flexibility allows the practitioner to experiment with different ***Cue*** approaches to aid each child towards his best performance in emotion recognition.

Several reinforcements are possible each time the child selects the appropriate doll. The guide can be cued to say "*THAT'S GOOD, THAT'S <emotion>*" when the correct choice is selected. Additionally, the practitioner may opt to have the system play an entertaining reinforcement video clip, such as a Tigger song. Our objective was to offer as much flexibility to the practitioner as possible for customizing the screen interface for a particular session or specific child. This is especially important because autistic children often have unique idiosyncratic behaviours.

The Child interface consists of one or more elements set up by the practitioner as just discussed. Figure 1 shows the screen seen by the child, set up here to show the video clip in the middle, the emotion icon at top, the dwarf face at left, the label of the emotion at bottom, and the guide at right. Images of the icon, dwarf, word, and guide always appear in the same spot when they are selected.

The child interacts with the system through a plush toy interface. Four interactive dwarves provide a tangible interface to the system, so that the child does not have to use a keyboard or mouse. Images of the dwarves, representing angry, happy, surprise, and sad, are pictured in figure 2, just as they appear on the screen when selected for display. The dolls serve as engaging input devices to the system: they are fun to hold and add a playful flavour to the interaction.

Figure 2. Pictures of the Dwarves

The system design has two modes of interaction -- an applied behaviour mode and a story-based mode. The first mode displays short clips, one at a time, from various child program sources and the second mode displays an entire movie with the story segmented by the emotions. When the video freezes, the interaction is the same for both modes until the correct doll is selected.

Working with researchers at the Dan Marino Center, Ft. Lauderdale, Florida, we designed the system to track performance information requested by the therapists. For each session, the system would record the child profiles, the system configuration, clip configuration, and response times for each interaction. *Response Rate* was computed to track the number of training trials (this measure includes both correct and incorrect responses by the child, normalised by the time of the session), *Accuracy* was computed as an index of how effective the training procedures are for teaching the children to match the targeted emotion: it consists of a ratio of correct matches over total attempted matches, for each trial. Finally, *Fluency* was computed

and used as a performance summary of how many correct responses were made – it is a measure that combines response rate and accuracy. An accompanying thesis (Blocher 1999) provides formulas for these measures.

3. EVALUATION & RESULTS

A pilot study was conducted to determine whether ASQ was engaging to children with autism and whether this type of an application may potentially help children learn emotion recognition. Children diagnosed within the autistic spectrum served as subjects. Subjects were recruited as volunteers through advertisements posted at the Dan Marino Child NETT Center in Fort Lauderdale. Standardised assessment tools, as well as direct observation by trained psychologists and neurologists, were used to identify children whose primary deficits are related to social-emotional responding and appropriate affect.

To participate in the pilot study, children needed to come to the center to play with ASQ for at least three days of sessions, each day's session lasting up to one hour. Nineteen different children with deficits along the pervasive development disorder (PDD) or autism spectrum from the Dan Marino Center were exposed to ASQ. Six of these nineteen children were observed over three days.

The study was conducted at the Dan Marino Center in one of their therapy rooms. The room was eight by eight feet, with one outside window and one window to another office. A laptop ran the ASQ application. The four dwarf dolls were the child's input devices to the application. Each toy doll was loosely positioned on the table on a reclining board adhered to the table with Velcro pads. The dolls could be picked up easily by the child, but were intended to remain on their stand because it was found to be easier for the child to press the belt-buckle of the chosen doll when the doll was on a hard surface (Figure 3).

Figure 3. Child Testing

The goal was to see if children can correctly match the emotion presented on the child-screen to the emotion represented by each doll. For experimental control the same dolls were used with each child, and all children were tested with the applied behaviour mode (vs. story mode). The automated training was arranged to teach children to "match" four different emotion expressions: happy, sad, angry, and surprised. A standard discrete-trial training procedure with the automated

application was used. Subjects sat facing the child-screen that exhibited specific emotional expressions under appropriate contexts within the child's immediate visual field. A video clip played for between 1 and 30 seconds. The clip displayed a scene in which an emotion was expressed by a character on the screen. The screen 'froze' on the emotional expression and waited for the child to touch the doll with the matching emotional expression (correct doll). After a pre-set time elapsed, a specific sequence of visual prompts displayed on the computer monitor and auditory prompts played through the computer speakers.

If the child touched the doll with the corresponding emotional expression (correct doll), then the system provided an online guide that audibly stated "Good, That's <correct emotion selected>," and an optional playful clip started to play on the child-screen. The application then displayed another clip depicting emotional content randomly pulled from the application.

If the child did not select a doll or if he selected the incorrect (non-matching) doll, the online guide provided a verbal prompt: "Match <correct emotion>" for no doll selection, or "That's <incorrect emotion>, Match <correct emotion>" for incorrect doll selection. The system waited for a set time configured by the practitioner and repeated its prompts until the child selected the correct doll. An optional replay of the clip could be set up before the session, in which case the application replays that same clip and proceeds with the specified order of prompts configured in the set up. If the child still fails to select the correct doll, the practitioner assists the child and repeats the verbal prompt and provides a physical prompt, e.g., pointing to the correct doll. If the child selects the correct doll but doesn't touch the doll after the physical prompt is provided, then physical assistance is given to insure that the child touches the correct doll. This procedure was used to replicate the discrete trials.

Two low functioning autistic children, between the ages of 2 and 3, engaged in the video clips yet displayed little interest in the doll interface without direct assistance. One boy, age 4, demonstrated an understanding of the interaction, although he struggled to match the appropriate doll. Another boy, aged 5, appeared to understand the interaction, yet had such a soft touch that he required assistance in touching the doll so that the system could detect what was selected.

A three-year-old child, with Spanish as the native tongue, appeared very interested in the application regardless of the language difference. He and his family were visiting from Chili and together they played with ASQ for one hour. Earlier two visiting neurologists from Argentina sat in on the child session and they were certain that the screen interface had too many images (referring to the icon, word, and dwarf's face) and thought that the dolls were not a good interface. After they saw this boy interact with the application, both the physicians and the boy's parents were surprised at this boy's quick adaptation to the doll interface and his ability to recognize the emotions, despite the screen interface.

As suspected, higher functioning and older children, age 6-9, demonstrated ease with understanding the doll interaction, exhibited pleasure with the gaming aspect, and needed few of the helpful screen cues to make their selection. They were able to match emotional expressions displayed on their screen by selecting the correct doll after only a few sessions. One boy mimicked the displayed emotions on the screen.

His mother reported that he was able to recognize other people's emotional expressions at home also.

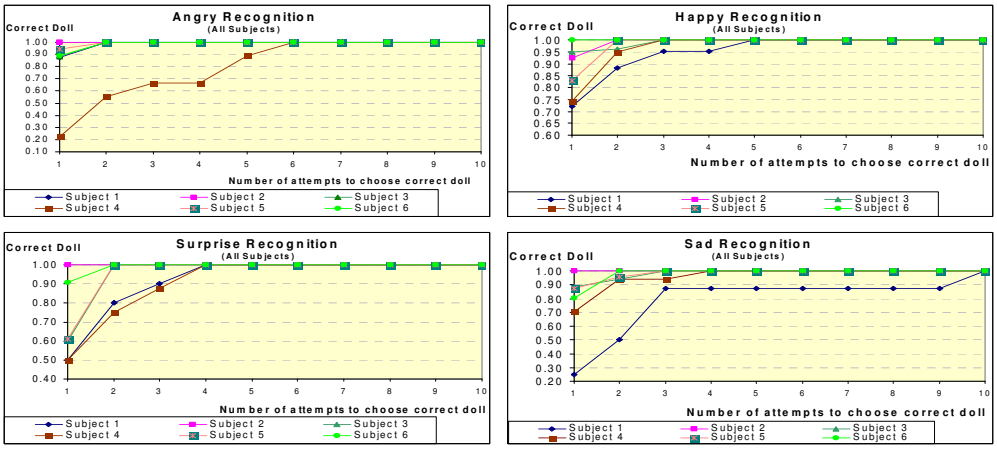


Figure 4. Recognition results for six kids (note some of the vertical scales differ.)

4. CONCLUSION

ASQ was considered successful at being engaging based on comments from the staff at Dan Marino as well as parent feedback and observations by the child study practitioner. Furthermore, the statistical findings suggest that emotion matching occurred in most child cases, with some children showing improvements in their performance at matching emotions over three sessions. Figure 4 shows combined recognition results for the six kids – where higher curves indicate better recognition rates. For example, when an angry clip was played, all kids (except subject 4) were able to correctly pick the angry doll when given two attempts. Subject 4 took many more tries than necessary to select the angry doll. For most of the kids, anger was the emotion most reliably recognised, while surprise and sadness were harder to get right. Five of the six kids were able to match many of the emotions on the first day. A three-year-old child showed results that he could recognise more samples of an emotion with each additional session of interaction. What the data did not provide is conclusive evidence that ASQ taught emotion recognition: it is possible that the children's performance improvement was due to something besides emotion recognition. A study including base line tests before and after using the system over a longer duration would present results that are more conclusive.

Although the ASQ system can measure improvements by a child while using the system, it does not know about improvements the child may show in real life,

outside the computer world. To understand how the effects of the system generalise to real life, long-term follow-up is needed. One mother did report that her son said, "I'm happy" with a smile on his face at the dinner table with the family. She doesn't remember him expressing himself like that before. Also, she said that when he was picked up from school he asked if he could go play with the dwarves.

Overall, the technology was found to be interesting and engaging to the children; the nineteen children interacting with ASQ were engaged in the application and seemingly enjoyed their interaction. Most children showed interest in the screen interface design and its multimedia elements. However, there were occasionally reliability problems with the wireless communication provided by the dolls; one practitioner suggested that it would be more reliable to use hard-wired buttons with the doll's faces printed on them. However, we observed in some sessions that the child and a parent would each hold a doll, and when the parent was holding the doll that the child needed to select, the child would turn to the parent to achieve this; thus, the physical nature of the interface actually aided in helping the child with eye contact and shared experiences referred to as joint-attention, another area where autistic kids tend to have difficulty. Thus, we continue to see promise in the playful doll interface, despite occasional reliability problems with this first system.

Although the ASQ system was developed for autistic children, it is our belief that all people can benefit from learning systems that enrich emotional understanding and positively affect our social quotient. For example, even professional physicians could benefit from improved abilities to recognise and respond to human emotion; indeed, lack of such abilities has been found to contribute to costly malpractice lawsuits.

It is noteworthy that many adult autistics like interacting with computers and with people over the Internet; some autistics have said that such computer-mediated interaction levels the playing field for them. Computers that inhibit affective communication, while a help to autistics, may be diminishing the opportunities of all of us to practice emotional communication. Increased time online, spent interacting through channels that limit transmission of emotional information may be costing all of us the opportunity to improve our affective social quotient. This project, while just a first step, shows that computers may also be able to serve as tools that help improve affective communication for all.

5. BIBLIOGRAPHY

Baron-Cohen, S., Volkmar, F Ed. *Handbook of Autism and Pervasive Developmental Disorders*, John Wiley & Sons, NY. 1997.

Blocher, K. Affective Social Quest: Teaching Emotion Recognition with Interactive Media & Wireless Expressive Toys, SM Thesis, MIT 1999.

Eckman, P. *An Argument for Basic Emotions*. Cognition and Emotion, Vol 6, 1992.

Lubin, D. *Behaviour Engineering through Computer-based Automation: A system for training free-hand drafting skills*. Auburn University PhD. Thesis. 1989.

Mayer-Johnson web page http://www.mayerjohnson.com/ 1999.

Picard R. *Affective Computing*. MIT Press, Cambridge, MA, 1997.

Sigman M. & Capps L. *Children with Autism : A Developmental Perspective*. Harvard Press, Cambridge, MA 1997.

Tuchman, R. Conversation while conducting experiments at Dan Marino Center September 1999.