Rover@Home: Computer Mediated Remote Interaction Between Humans and Dogs

Benjamin Ishak Resner

B.A. Physics, Cornell University, 1990

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning in Partial Fulfillment for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology

September 2001

Author
Benjamin Ishak Resner
Program in Media Arts and Sciences
August 10, 2001

Certified by
Dr. Bruce Blumberg
Associate Professor of Media Arts & Sciences

Accepted by
Dr. Andrew Lippman
Chair, Department Committee on Graduate Studies
Program in Media Arts and Sciences

Rover@Home: Computer Mediated Remote Interaction Between Humans and Dogs

Benjamin Ishak Resner

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning in partial fulfillment of the requirements for the degree of Master of Science at the Massachusetts Institute of Technology.

Abstract

In this thesis we create a method to allow dogs and humans to interact over the Internet. In particular, we generalize an established dog training technique known as "clicker-training" such that the remote and co-located interactions are reported by dog owners to be similar. In the process of creating this computer-mediated interaction, we learn what it means to design an interface for a creature with very different sensory modalities than humans. Dogs are not "furry humans" but entirely different creatures with very different perceptual, motor, and cognitive systems than humans. This work is significant because by systematically applying HCI design principles to non-humans, we include animals in the HCI community. This creates an opportunity for the evaluation of the generality of much HCI literature, as well as increasing the sources from which we can draw inspiration.

Thesis Supervisor: Dr. Bruce Blumberg, Associate Professor of Media Arts & Sciences

Rover@Home: Computer Mediated Remote Interaction Between Humans and Dogs

Benjamin Ishak Resner

Thesis Advisor
Dr. Bruce Blumberg
Associate Professor of Media Arts & Sciences
MIT Media Laboratory

Thesis Reader
Dr. Hiroshi Ishii
Professor of Media Arts & Sciences
MIT Media Laboratory

Thesis Reader
Dr. Irene M. Pepperberg
Visiting Associate Professor at the MIT Media Laboratory
Associate Professor of Ecology & Evolutionary Biology
University of Arizona, Tucson

TABLE OF CONTENTS

1 Introduction

The goal of this thesis is to combine principles of interface design and animal behavior to create an interface to allow a dog to interact with a remote human via a computer.

This thesis begins by situating the study of animal-computer interaction in the context of the maturing field of human-computer interaction (HCI). With this relationship established we then search the HCI literature for specific methodologies that can be elegantly generalized to non-human animals. In particular, we discuss User Centered Design (UCD), a popular and well-established user interface methodology that turns out to have relevance beyond humans. We also evaluate some of the more modern descendants of UCD.

This work is equally informed by animal behavior. Some notable experiments in animal psychology are evaluated from the perspective of sound interface design. Some of the most significant research in animal communication and language can been seen as the result of carefully designed interfaces that take into account the uniqueness of the species being studied.

In order to build interface devices for dogs, an understanding of their physiology, psychology, and social behavior is necessary. We analyze canine sight, smell, hearing, and touch, paying attention to the technical feasibility of piping each sensory mode over a remote network. We also look at how these sensory modes could combine to build a sustaining remote social interaction.

Sensitivity to physiological and psychological differences between humans and dogs suggests computer-mediated interaction will have asymmetric interfaces for each creature – how the dog experiences the interaction is different from how the human experiences it. Because dogs and humans have very different input and output systems, interfaces for each creature will be correspondingly different.

With this theoretical framework in place, we describe the application of these principles to the construction of an artifact for supporting remote dog-human interaction. Ideally, this device will take advantage of established rituals. Established rituals provide a conceptual model for a remote interaction that will leverage familiar play patterns.

One such ritual is clicker training. This is not only an increasingly popular dog training method, but also its affordances are compatible with current computer input and output devices. Clicker training is based on applied operant conditioning, and uses a click sound in conjunction with food treat rewards as a powerful tool for shaping canine behavior. Perhaps of greatest importance is that pet owners perceive clicker training to be an entertaining activity for themselves and their pets. For this reason, along with technical feasibility, clicker-training was chosen as the model for a remote interaction.

The basic Rover@Home device consists of a computer-controlled treat dispenser to reward the dog, a webcam for visually monitoring the dog, and a speaker for the dog to hear clicks and the owner's voice. Engineering challenges include bandwidth limitations and latency. We designed novel solutions for these impediments to realtime remote interactions.

We tested the device with several dogs and present the results of two sets of field trails. In both instances, the dog – owner interaction encouraged further work. Based on these results, we present a more rigorous process for introducing dogs and owners to the Rover@Home setup. We also outline a future research direction as well as applications of Rover@Home.

A technical appendix describes the software and hardware developed for this thesis.

2 Background and Motivation

There are three main inspirations for this work. First is the emotional and egalitarian (and perhaps irrational) desire to create an interaction that allows dogs and humans to sustain a meaningful interaction while separated. Second is the intellectual opportunity to systematically apply HCI design principles to non-humans. Finally, we see this work as being part of a larger trend of affording dogs many of the same entitlements as humans.

2.1 Emotional Need for Remote Dog-Human Interaction

Thirty-three percent of dog-owning Americans talk to their dogs on the phone or through an answering machine when away from home (Consumer Reports, 1998). This demonstrated human desire for remote interaction is at the core of Rover@Home. There is clearly a craving to remain in touch with pets while separated. Live webcams are increasingly becoming a standard part of "doggie day care" so remote owners can verify that their dogs are happy and well cared for[1]. We have received anecdotal reports of owners who take employment only close to home so they can visit their dog during lunch.

This desire for pet closeness should not be surprising in light of the deep emotional bonds humans form with their pets. "He is my firstborn" declares a pet owner, without any sense of insincerity. In the United States, 65% of dog owners report buying a Christmas gift for their dog, and 21% report celebrating their dog's birthday. 41% of Americans display a picture of their pet in their home and the most popular place for a dog to sleep is on the owner's bed (American Pet Association, 1998). 87% of dog owners consider their dog to be a full-fledged family member. Perhaps

[1] Three easily found (and cleverly named) examples are: "Canine To Five", Lake Worth, FL (http://www.caninetofive.com/), "It's a Ruff Life", Phoenix, AZ (http://www.itsarufflife.com/), and "Central Bark", Los Angeles, CA (http://www.lacentralbark.com/).

most significant is that one third of dog owners report they are closer to their dog than to any family member (Barker and Barker, 1988).

From this data, as well as informal observations of humans with their dogs, we conclude there is a need for a means for humans and dogs to interact while separated. As the world becomes increasingly wired to allow more remote interactions between humans through telephone, email, videoconferencing, online chat, Internet telephony, and fax, it makes sense to research a method to allow remote connections with the furry members of one's family.

2.2 Intellectual Merit of Applying HCI Design to Non-Humans

Frameworks for developing computer interfaces for non-human creatures simply do not exist. If we want the process of creating a remote dog-human interaction to be guided by some type of design principle, we are either going to have to start from first principles or to generalize an existing body of work to our needs. We find the field of Human Computer Interaction (HCI) particularly relevant because it deals with the adaptation of mechanized systems to biological systems. Insofar as humans and dogs both have brains, eyes, and ears, but not hard disks, microprocessors, and keyboards, the interaction between dogs and humans could likely be guided by our understanding of successful interactions between humans and computers.

To the best of our knowledge, this is the first instance of non-human animals being studied under the framework of HCI. Computer interfaces for animals, most notably primates, have been researched and developed, but none of them has made reference to HCI literature (Rumbaugh, 1977; Gardner, Gardner et al., 1989; Savage-Rumbaugh and Lewin, 1994; Fouts, 1997). Utilizing HCI for the purposes of human-animal communication creates the opportunity to evaluate the generality of this body of work to non-human animals. Even the most anti-canine would probably find merit in the prospect of discovering unexpected universality to a body of research.

In addition to using HCI research to guide development of computer-based interactions for dogs, one can take the results of these investigations and bring them back into the body of HCI research. We will show that including dogs under the HCI umbrella is not only good for dogs, but good for humans as well.

2.3 Societal Trend of Affording Dogs Similar Entitlements as Humans

While Rover@Home may seem far-fetched (no pun intended), it appropriately fits into larger trends of affording pets many of the same entitlements as humans. This is becoming especially true in modern countries where couples have fewer children later in life. Humans have been living in the company of domesticated dogs for between 10,000 and 100,000 years (Clutton-Brock, 1995), inviting them into their homes, and often treating them as furry humans or surrogate children. Including dogs as beneficiaries of computer technology is a logical continuation of this trend.

Americans certainly spend freely on their pets. The over 43 million dogs in the United States representing ownership by 60% of all US households (American Pet Association, 1998) consume $30 billion a year on pet food, supplies, veterinary care, and other services (Tribune Business News, 2000). This breaks down to an average of $650 per dog, per year.

There is certainly no shortage of opportunity for pet related spending. Available to the discerning customer are pet health food (Duin, 1998), pet nutritionists, pet aromatherapists, pet acupuncturists, pet loss support groups (Williams, 2001), pet psychiatry, pet massage therapy, and pet chauffeurs (Iovine, 2000) to transport a beloved pet between play dates. There is even a minister in California who performs wedding ceremonies for dogs about to be mated (Beck and Katcher, 1996).

Figure 2-1: Dog sporting a $150 leather jacket from Coach. Coach is an upscale human clothing accessory retailer that sells quality canine clothing without any touch of irony.

© 2001 Coach http://www.coach.com

Associated with this trend is a strong anthropomorphization of our pets. The place people would leave their pets while on vacation used to be called a "kennel." Now these establishments typically are named "Pet Spas," "Pet Camp," "Pet Daycare," "Pet School," and "Pet Retreats." All of these words connote a place a human would gladly visit. In contrast, "kennel" conjures images of caged dogs given little more than food and water.

The adaptation of human products and services to dogs is not limited to frivolous or commercial endeavors. Serious science is also pursuing dog-related endeavors. There is a dog genome project, (Rine, 2001) attempting to map dog DNA in much the same way as the human genome project is trying to chart the human genome. Similarly, the "Visible Animal Project" (Böttcher, 1999) has as its first creature a cross-sectioned dog available for online viewing.

Pets are also gaining legal status. In addition to several books advocating "personhood" for animals[2], the Berkeley city council has become the third city in the United States to pass a resolution requiring all legal language relating to human-animal relationships to refer to the human as the pet's "guardian" as opposed to "owner" (Brown, 2001). The city felt the word "owner" implied that a pet was enslaved or otherwise robbed of its status as an independent creature. The city attorney has been advised that all pending civic activity should reflect this change[3].

Six states have outlawed dogs in the back of pickup trucks, signifying the first legislative activity centered on pet automotive safety. Similarly, Saab has declared itself to be the first automotive company to come out with a line of products specifically aimed at pet safety and comfort in cars. Properly restrained dogs are more than a personal safety decision. Unrestrained dogs involved in a collision hamper rescue operations and create additional hazards, which presents a public-safety issue that will likely drive additional legislation and consumer demand for associated pet safety products.

3 Design Goals and Methodology

Here we describe the intellectual process for understanding how to build a device leading to computer-mediated interaction between dogs and humans. This discussion starts with an analysis of how the body of existing HCI work is relevant to the study of computer interfaces for animals. Care has been taken not to over-generalize to all animal life, but certainly many of the lessons here will likely be applicable to creatures other than domesticated dogs.

3.1 Situating Animal-Computer Interactions in an HCI Framework.

In many ways, the inquiry into animal-computer interaction is well underway. If humans are considered to be a subset of animals, the vast and growing body of human-computer interaction can be seen as being a special case of the wider field of animal-computer interaction. Rather than put all humans into one category and lump all other non-human animals into another category, it is preferable to see humans as a subclass of all animals.

---

[2] For example, "Rattling the Cage" by Steve Wise (Perseus Books, 2001), "The Case for Animal Rights" by Tom Regan (Univ California Press, 1985), and "Animal Liberation" by Peter Singer (Avon/Hearst Corporation, 1991)

[3] This thesis makes use of the word "owner" to describe the person who feeds and shelters his or her dog. Similarly, we use word "pet" as opposed to "companion animal".

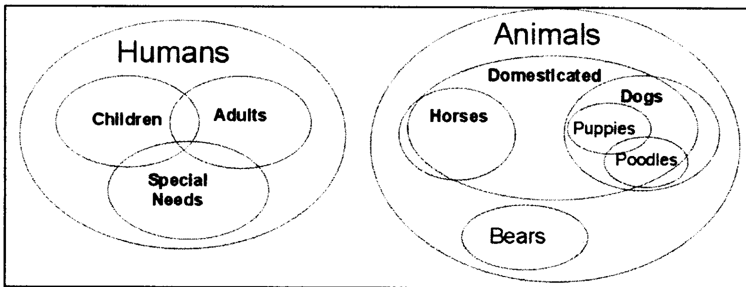


Figure 3-1a: A possible framework for investigating animal-computer interactions. Humans and Animals are distinct groups, each with their own field of study.

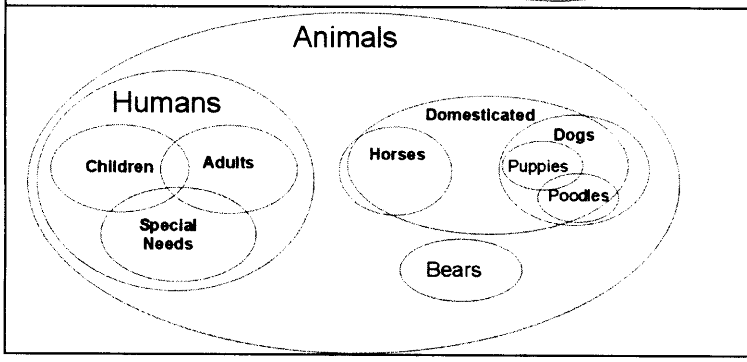


Figure 3-1b: Our approach to investigating animal-computer interactions. Humans are simply the most studied member of a larger category. Domesticated dogs are also members of this category.

This distinction highlights two important points. First, it makes the existing body of human-computer interactions much more relevant to our study. If human-computer knowledge is looked at as simply a special case of animal-computer interaction, one has a much easier time generalizing its methodologies and lessons to other animals. The body of HCI literature immediately becomes more relevant to this inquiry when humans and animals are not separated into two distinct categories.

Similarly, it will be easier to apply insights about non-human animals back to humans when humans are seen as belonging to the same superclass as domesticated dogs. A secondary hope of this thesis is to increase understanding of human-computer interactions through the study of animals. Similar to how many scientific endeavors have greatly benefited from animal experimentation, human-computer interaction might be a similar beneficiary[4].

Second, this grouping highlights the fallacy of bundling all non-human animals into one big category and studying them en-masse. Just as humans have been studied as a special case of animals, all other species must also be given individual attention. Realistically, as more species are studied, we would hope to gain greater appreciation for general rules of animal-computer interaction, but as an initial approach, it would be a mistake to study horses, bears, and dogs in the same iteration. For these reasons, we have chosen to view dogs as having the same categorical relationship to humans as they do to bears.

We use liberally use the term "HCI" to include subspecialties perhaps more appropriate to the artifacts developed in this thesis. Technically, HCI refers only to actual interaction between humans and computers, whereas Rover@Home is

[4] Fortunately, animal based HCI research would not typically come at the expense of the animal's welfare.

technology mediated interaction between humans and animals. Therefore, Rover@Home more appropriately belongs in the realm of computer supported cooperative work (CSCW) and human-computer, computer-human (HC-CH) interaction. Because both these specialties include HCI as a subcomponent, we generalize the literature under the all-purpose term "HCI." Future iterations of this work would do well to be more specific about identifying relevant features of the various research agendas that investigate how humans and technology interact, and how technology in general and computers in particular support human-human interaction.

3.2 Scope

The focus of this research is stated to be domesticated dogs. But as the previous section might suggest, it should not be taken for granted that this grouping is optimal. Just as we claim the category of "non-human animals" is too general, so too might be "dogs." There is the option to specialize further within dogs and examine only certain breeds. On the other hand, the category of "dogs" might be overly specific.

The decision to study domesticated dogs is somewhat arbitrary, made in part for research convenience – we have access to domesticated dogs as research subjects, but not necessarily significant numbers of any particular breed. There also seems to be an intuitive logic in this grouping. People often define themselves as "cat people" or "dog people," but less typically as "herding dog people" versus "retriever dog people."

Domesticated dogs can be further subdivided along breed, breed group (e.g. terrier, sheepdog, retrievers), task (scenthounds, companion, herding, hunting)[5], gender, or even size. Dogs are noted for the remarkable plasticity they exhibit across breeds, with wide variations in stamina, sensory perception, size, agility, and emotion (Fogle, 1990) But these are all matters of degree. No breed possesses an ability completely absent in another breed. Furthermore, one of the most similar features between dogs is their cognitive power (Fogle, 1990). And because the artifacts of this research rely less on any specific motor function and more on a dog's cognitive power, the grouping of dogs makes sense.

To some degree this may seem like using the conclusion to justify the hypothesis, and perhaps it is. Had the interaction developed depended on a dog's instinctive herding impulse or weighing a certain size, we would certainly have had to take note of breed

[5] There is no shortage of means to group breeds. The American Kennel Club classifies dogs into "Sporting", "Hounds", "Working", "Terriers", "Toys", "Non-Sporting", "Herding", and "Miscellaneous". The FCI (Federation Cynologique Internationale) groups dogs according to: "Sheepdogs and Cattledogs", "Pinschers, Schnauzers, mastiffs, and Swiss mountain & cattledogs", "Terriers", "Dachshunds", "Primiitve Type Dogs and Spitzes", "Scenthounds and related breeds", "Pointers", "Retrievers, Water Dogs, and Flushing Dogs", "Companions and Toys", and "Sighthounds".

differentiation. But because we ultimately chose an interaction based on a dog's cognitive ability, breed differences count for less.

Finally, one of the goals of this research is to develop an interaction for all domestic dogs, not just a subset. There was no desire to give one type of dog preferential treatment.

Future interactions based on this research would do well to pay attention to breed differences, and future evaluations of this interaction would definitely benefit from an analysis of the role of breed and breed group. We leave this for future work.

3.3 HCI Background applied to Animal-Computer Interaction Design

The human-computer literature is replete with methodologies for creating human-usable machine interfaces, and if humans are understood to be a subset of animals, all these methodologies are potentially generalizeable to all animals. It is our hope to extract particularly relevant features of a few of these methodologies and develop a design methodology appropriate for development of technology for dogs. We will then attempt to understand how this relates back to the more specific field of human-computer interactions.

We start with Donald Norman's work on User Centered Design (UCD). This is widely regarded as a basis of much modern interface work. This work is chosen not just for its popularity but also because it contains design sophistication without human-specific overspecialization. After a discussion of UCD, we proceed to look at three post-UCD paradigms.

A first pass for much of our approach is to read the HCI literature and substitute the word "pet" or "dog" for the word "human" or "child." In many instances, this substitution works well, and yields useful and functional insights. In other cases, such substitutions are inappropriate and even comical. In either case, it is safe to assume the authors never expected their work to be applied outside the domain of humans, so any relevance can readily be ascribed to true universality and not an academic desire to create one big theory that explains everything.

3.3.1 Prelude to UCD: Animal Trainers and Interface Designers

Much of User Centered Design is an effort to bridge the gap between the designer's intentions and the user's desires (Norman and Draper, 1986). All too often the user cannot understand the interface, and the designer cannot understand why the user is confused. To the designer, the interface is perfectly logical and needs no explanation. To the user, the interface is obscure, and intuitive leaps from one feature to the next are impossible. Ideally all users would have continuous access to the designers of their technology, but this is not realistic. Therefore, communicating meaning and coherence across this gap is the goal of interface designers. A well-implemented interface harmonizes the user with the designer, allowing the user to operate the software as the designer intended (Norman and Draper, 1986).

Overcoming this communication gulf is also the goal of animal trainers. Trainers seek to shape a desired behavior, or extinguish destructive and anti-social behaviors (Ramirez, 1999). Yet the channels of communication between trainer and animal are similarly circuitous. Trainers cannot use plain English (or French, or Russian) to explain to the animal what they want, and why it is important. Trainers can become frustrated trying to understand why the animal has difficulty acquiring a new behavior or extinguishing an objectionable one. Similarly, animals can exhibit heightened distractibility during these difficult times. Both animal trainers and interface designers rely on a detailed knowledge of their target audience to nuance the interaction in ways that optimize clarity and information transfer.

But while interface designers may become frustrated with users, there is little controversy about the common ground they do share. Interface designers can assume a color or shape placed on a computer screen will appear as that color or shape to the user. Similarly, interface designers can intuit user keyboard and mouse dexterity by extrapolating from their own experiences. Interface designers for other humans are correct to assume their users bear a certain amount of physiological and cognitive similarity to themselves.

This becomes less true as the target user deviates from the mainstream. Disabled and young users have a spectrum of special needs and different abilities that may not be readily inferred from one's own experience. The younger or more disabled the user, the more these basic mechanisms of acquiring and processing information and acting upon it will be different from those of the typical interface designer. Furthermore, it is not necessarily the case that this user population has overall diminished capabilities. Visually impaired humans become more sensitive to taste and smell (Smith, Doty et al., 1993), and newborn human babies prioritize stimuli differently from adults. An optimized interface would take advantage of these heightened abilities and different attention to stimuli.

This assumption of sensory and behavioral correspondence is even less true between trainers and animals. Animals have very different perceptual, cognitive, and motor abilities from humans, some of which are below human abilities while others are quite superior. But much less can be extrapolated from our own experiences. An interface designer focusing on animals would do well to take advantage of some of the superior sensory abilities many non-human animals posses. Studying animal anatomy and behavior in order to adopt a "least-common-denominator" approach to interface design for human-animal interaction is only slightly better than ignoring animal uniqueness entirely.

Given that User Centered Design is intended to be used by humans for humans, it is not surprising that a thorough understanding of the perceptual and motor system is not emphasized. Much attention is paid to understanding the cognitive and motivational model of the user, but not to basic sensory abilities (Norman and Draper, 1986). However, it is definitely within the spirit of User Centered Design to implement this step as appropriate. To this end, an analysis of canine physiology with respect to currently available technology is included.

3.3.2 UCD and Non-Human Animals

With consideration given to the importance of understanding non-human animal input and output modes, we are ready to look at four intellectual cornerstones of UCD. Below is a discussion of task domains, affordances, cognitive modeling, and direct manipulation, and how they apply to non-human animals.

The following four sections are all taken from "User Centered Design," edited by Donald Norman and Stephen Draper (Norman and Draper, 1986).

3.3.2.1 Task Domains

Over-generality sinks interfaces. The more an interaction can address a specific need, the greater its chances for success. In fact, the successful identification of a human task can compensate for a poorly designed interface. If a computer allows a human to perform an action a he is motivated to perform he will be willing to learn and use interfaces that are difficult to understand and awkward to use. Early spreadsheets such as VISICALC were popular not because of a wonderful interface, but because they allowed accountants to model financial interactions that were previously painfully tedious. The task of rapid scenario manipulation compensated for the counter-intuitive keyboard-only interface of VISICALC.

Difficulty in learning a new interface should ideally be because of poor understanding of the task domain, not poor understanding of the interface domain. If someone is having difficulty with a drawing program, it is hopefully because the user has an incomplete understanding of color mixing or airbrush operation. Even the best-designed interface cannot compensate for lack of competency with the basic skills the interface represents. The most sophisticated word processor, complete with spell and grammar checkers, cannot compose compelling publishable fiction. That struggle is left to the user.

Appropriate task domains are relevant to design for animals. One must be clear about what animal behavior is being replicated, augmented, or leveraged. The interface must allow transparent access to the task the animal seeks. The result of the technology interaction must be something the animal craves, and the path to that interaction must not be obscured by an opaque interface. Furthermore, if the task domain is clear, shortcomings in the interface will be compensated for by the desire to accomplish the given task.

3.3.2.2 Affordances

An interface should reveal its functionality through its affordances. A hammer comes with no instruction manual because it is clear what part is for hammering and what part is for holding. An interface should invite the user to interact in the intended manner. If someone is given a hammer for the first time in their life (such as a child), they may attempt to hold it incorrectly, but experimentation will quickly lead to the optimal grip the hammer designer intended.

As interfaces become more complex and control greater amounts of functionality, simple single-dimensional hand-tool like affordances become more complex. But the idea of keeping interfaces self-discoverable through exploration remains intact. Affordances emphasis "recognition over recall," where an object's use does not have to be memorized, but can be readily inferred from its design.

Affordances are especially relevant to animals because, like many humans, animals do not read manuals. Artifacts of an animal-computer interaction must present themselves in an intuitive manner that requires minimal training. Although it can be expected animals will learn or be trained to participate in an interaction with the computer, careful attention to affordances will minimize this step and make it more natural. Affordances consistent with an animal's natural biology also make learning new features or nuances much easier.

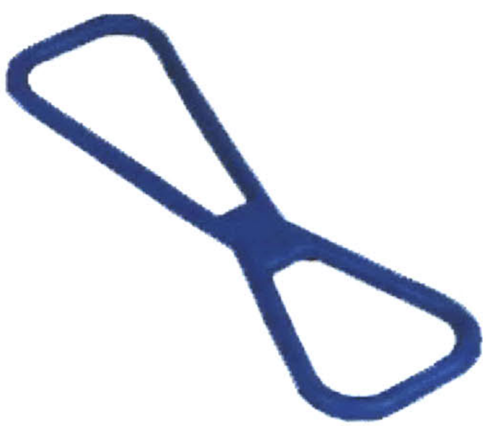


Figure 3-2: Dog-human tug toy. The shape of the toy implies where it should be grabbed and what it should be used for. It's a fortuitous coincidence the affordance for canine gripping in mouth is the same as human gripping in hand. Neither dog nor human has to be trained to use this toy. The proper use is implicit in the design

3.3.2.3 Cognitive Modeling

An interface should allow a user to interact with his or her environment in a manner consistent with his or her cognitive model of the task. If I am using a word processor, and want to cut and paste, I create a cognitive model of removing some text from the document and placing it into a temporary container, followed by a model of inserting the contents of that container into a different location in the document. A well-designed interface takes advantage of user's cognitive models by acting consistently with their expectations.

When applying this guideline to animals, we take "cognitive" to include all behavioral models, encompassing instinctual, conditioned, and learned. What is essential is the interaction must be based on some known behavior already resident in the animal. For example, foraging instincts can provide a cognitive / behavioral model when designing an interaction for birds. Similarly, it can be argued the success of first-person "shoot-em-up" games such as "Doom" are based on cultural and/or biological predator instincts in human males.

3.3.2.4 Direct Manipulation

In addition to how an interface should leverage a cognitive model, it should, to the extent possible, allow a user to feel as if he or she is directly interacting with the task elements. The user's interaction with the task should not be obscured by the interface, but be focused on the actual task. In the most literal sense, direct manipulation implies the computer is completely transparent and the user is physically moving objects with his or her hands.

This idea is especially important for animals. It is not reasonable to expect them to have a sophisticated understanding of symbolism or iconography. For them, interfaces need to be as literal and direct as possible. Dogs cannot be expected to understand that pressing a button is the same as interacting with its owner.

3.3.3 Post-UCD paradigms

User Centered Design has been successful as establishing a foundation for human-computer interaction development. Recent work has focused on how the user is studied and included in the design process. The general trend has been towards a more naturalistic approach to studying the user and his or her relationship with the computer. Augmenting laboratory studies of human motor control and perceptual abilities are methods inspired by fields such as Ethnography and Anthropology (Preece, 1994). These evaluation methods study the user's interaction in a natural social setting, typical of actual conditions under which the technology will be used. Although an understanding of the components is essential, we believe the evaluation of the complete social animal interacting with its environment is fundamental to the successful development of animal-computer interactions.

3.3.3.1 Contextual Inquiry

Contextual Inquiry (Beyer and Holtzblatt, 1997) tells us that researchers should collect data in the user's own environment as opposed to studying users in the laboratory. Much traditional HCI research focuses on human-factors such as Fitts Law, which is the mathematical expression of the time it takes to hit a target given the size of the target and the distance from the target (Fitts, 1954). Fitts Law does not take into account (nor does it need to) the environment in which the user resides. It is an expression of human physiological ability abstracted from any environmental concerns. However, psycho-motor investigations are only one part of the picture when designing effective user interfaces.

Although a similar functional understanding of canine cognitive and motor abilities is essential to building a successful interface for them, we agree that studying dogs in the full complexity of their natural world will yield insights beyond what a piecemeal deconstructionist approach can offer. This evolution towards interaction and collaboration over observation and analysis reflects a similar trend among some animal behavior researchers away from traditional Skinnerian analysis taking place in highly controlled laboratory studies, and towards an ethological approach that studies animals in their natural social and environmental settings (Gray, 1994). It is interesting to note that in the field of animal behavior, the more naturalistic the setting, the greater the reported results, and the more vociferous the criticism stemming from accusations of inadequate adherence to reproducible laboratory control (Fouts, 1997; Pepperberg, 1999).

3.3.3.2 Cooperative Inquiry

Continuing this line of reasoning, Allison Druin goes well beyond the laboratory study of humans interacting with computers and considers the children she works with to not be study subjects, but actual research participants with equal footing to the adult researchers. These "intergenerational" teams, as she calls them, form the core development group for the children's software she designs (Druin, 1999). She works in groups ranging in age from seven to adult, and the groups will stay together for as long as two years.

Is it fair to think of our interaction with pets as an "interspecies collaboration?" Can we truly consider dogs to be full-fledged research partners, deserving of having their names listed on peer-reviewed journal submissions? The children in Druin's teams range in age from 7 – 11 years old, and no mention is made of any cognitive or behavioral difficulties that might otherwise interfere with age-appropriate use of human language. Therefore, these children are able to express themselves verbally, draw pictures and diagrams representing internal ideas, and participate in collaborative activities with some understanding of adult capabilities.

Cooperative Inquiry is where the trend of increased user participation loses relevance to our application of HCI to dogs. Dogs clearly cannot communicate at the level of the children in Druin's teams. Although the trend of increased user participation is helpful in understanding dogs, Cooperative Inquiry is perhaps where we draw the line.

3.3.3.3 Anyone, Anywhere … Any Species

The slogan for the 2001 Conference Computer-Human Interaction (CHI) was "Anyone Anywhere", projecting a conference theme of designing for computer users who are traditionally left outside the technology umbrella. In his plenary address at

CHI 2001, Gregg C. Vanderheiden encouraged interface designers to think of disabled users even when designing for a mainstream audience (Vanderheiden, 2001). He cites the classic example of curbcuts – a utility designed for people in wheelchairs but employed by a much wider population such as bicyclists, parents with strollers, and delivery people. By thinking about a differently abled audience, the creators of curbcuts intentionally or unintentionally benefited a much wider population. Similarly, designing for people who are hard of hearing or low vision has utility for users who are in noisy or poorly lit environments. Expanding products to include these less mainstream users not only serves a valuable social good, but also increases their utility to the mainstream. Thinking of disabled audiences is a valuable design exercise in a product development cycle.

Extending this theme of "Anyone, Anywhere" to nonhuman animals is a logical next step. Thinking through how an animal would interact with a given device or interface might be a useful exercise in understanding various scenarios. In the "Future Work" section, we describe how developing a device to allow remote dog-human interaction has clarified our thinking on how a similar device would work for pre-verbal human children.

3.4 Previous Work in Animal – Computer Interaction

Although we believe this to the first work in computer-animal interaction to be explicitly placed within a HCI framework, prior relevant work is to be found in the field of animal behavior. Some of this work even includes use of a computing device as an intermediary between humans and animals. Whether a computer is used or not, much human understanding of animal behavior relies on an understanding of the "interface" between human and animal. This interface analysis is relevant to our inquiry.

3.4.1 Interface for Animal Communication

Successful language work with animals starts with an understanding of their "interface" with humans. Early work in animal communication labeled animals "less intelligent" because they could not talk with humans using human words. When Winthrop and Luella Kellogg raised a chimp, Gua, alongside their newborn son Donald, the hope was that by cross-fostering Gua in the same environment as a human child, Gua would pick up similar linguistic abilities. Unfortunately, Donald learned more from Gua that the other way around. The study concluded that chimpanzees physically develop much faster than humans, but are unable to learn human language. When Keith and Cathy Hayes repeated a similar experiment in 1952 with Viki, another cross-fostered chimp baby, the results were only slightly better. Vicki learned to say: "mama," "papa," "cup," and "up," but not much else. The conclusion was similar: chimpanzees were unable to learn human language (Gardner, Gardner et al., 1989).

It was not until 1969, when Beatrice and Allen Gardner raised Washoe with American Sign Language (ASL) as her primary means of communication, that scientists had any success teaching language to a primate. What the Gardners showed is that the barrier was not cognitive, but one of vocal production. Once Washoe was given a means to express herself, she could use her signs in novel contexts, for instance signing "water bird" for swan (Fouts, 1997). As soon as an appropriate interface was developed, Washoe was no longer constrained by the affordances of vocal-centric language, and could express herself closer to her cognitive capacity. Sign language took the language problem out of the interface domain and into the task domain. Previous research had suffered from an inability to get away from overly human-centered interaction design, and therefore only focused on a struggle with a poor interface.

In deciding to use sign language, the Gardners were influenced by research indicating the primary means for communication among wild chimpanzees was gestural. Although chimpanzees possess a small repertoire of fixed vocalizations, most of these were tied to the limbic system, a low-level area of the brain triggered by primal emotions such as fear or excitement (Fouts, 1997). By studying the physiology and social behavior of chimpanzees, the Gardners were able to develop an appropriate interface for human-animal communication.

Although the nature and degree of language acquisition in these primate studies have come under scrutiny, the breakthrough of relying on sign language instead of vocal production remains intact. The controversy revolves around whether or not the animals truly understand what they are signing, or are simply repeating learned patterns. But nobody debates that for chimpanzee language studies sign language is superior to vocal language (Shettleworth, 1998).

With an understanding that the human-animal interface is fundamental, researchers are able to better appreciate that non-human animals each have unique ways of interacting with the world. In this light, bird songs and dolphin chirps are not only pleasant noises, but unique modes of interaction. Recent work with Elephants shows seismic communication over extremely long distances (O'Connell-Rodwell, Arnason et al., 2001). Only with an appreciation for the sensitivity of an elephant's foot, can one harness this modality as a potential source of interaction.

These breakthroughs in human-animal communication are possible because of the careful evaluation of what the target animals are capable of doing. As long as researchers persisted in using a human-centered interface to communicate with animals, they were not set up for success. With an understanding of how animals interact with the world, these pioneers were able to break through the physiological barriers to expose the deeper cognitive mechanisms.

Although this language work has focused on direct human-animal communication, careful attention to the notion of interface has proven valuable. Below are listed three projects that include explicit interaction with machines.

3.4.2 Duane Rumbaugh and Lana

In an effort to reduce some of the ambiguity and subjectiveness of ASL interpretation, Duane Rumbaugh and colleagues created a custom computerized system for their chimpanzee Lana (Rumbaugh, 1977). All communication with Lana was through a special language, Yerkish, developed specifically for human-animal communication. By employing a computer instead of ASL, not only was a large source of ambiguity removed, but Lana's entire linguistic output could be recorded and stored for evaluation by impartial scientists not involved in the project. The goal was to remove accusations of cuing and imitation being mistaken for cognition and innovation in primate language studies. By having a computer mediate the entire interaction, all parties have equal access to the same data. Furthermore, because Lana was interacting with a computer, she could engage in training 24hours / day, not just when trainers were present.

Communication took place at a keyboard containing up to 4 sets of 25 "yerkish" lexigriaphical symbols. By creating grammatically correct sentences, Lana could request objects or social interaction. For instance "please machine give Lana coke" would dispense a Coca-Cola™ from a vending machine.

When Lana pressed a key, it would illuminate, giving her feedback as to her action. The lexigraph would also be projected to an "input" area above the keyboard, where she could view the sentence being formed. Pressing the "period" lexigraph ended the sentence and submitted the request to the computer. A separate "output" area projected responses from trainers or interrogations from trainers to which Lana could respond. Although social interaction and play with trainers to whom Lana had formed attachments was a powerful motivator for Lana, all verbal communication took place through the keyboard and screen

In this instance, the computer was introduced more to address the issue of credibility and control, than to enable an interaction not previously possible. In fact criticisms of this project center on how use of a keyboard interferes with normal social use of language. The necessity of the research animal being in front of the keyboard in order to communicate interferes with the animal's ability to improvise and perform word-play (Fouts, 1997).

Despite these criticisms, this is probably the best-known example of a non-human using a computer to communicate with a human.

3.4.3 Vivarium Program

The Vivarium Program was an ambitious effort by a group at Apple Computer to, among other things, create a computer interface for captive Gorillas. In particular, they were working with Koko, a Western Lowland Gorilla (Gorilla Foundation, 2001) with an American Sign Language vocabulary of 600 words. The ultimate goal of the project was the adaptation of a Macintosh II to allow Koko "to have a voice." Using an icon driven touch screen, Koko could interact with her world by causing the computer to emit human synthesized speech. Configuration screens available to the human researchers allow them to customize the screen to best suit Koko's needs (Clark, Ferrara et al., 1990).

Their approach seems to mirror the successful teaching of American Sign Language to Koko, except instead of communicating with hand gestures, she communicates via a computer attached to a voice synthesizer.

A computer interface, however, goes beyond the exclusively communicative ability of ASL by allowing Koko to control her environment. By attaching various controls to the computer, Koko would have the ability to alter the world around her. One of their investigatory questions was whether or not Koko would begin to do for herself tasks she previously relied on others to implement. Unfortunately, the project has not published results much beyond the concept phase.

3.4.4 Dolphin and Keyboard Research

A research program at Marine World Africa USA constructed a special keyboard for dolphins, which was used to by the dolphins to associate vocal labels with objects (Reiss and McCowan, 1993). The dolphins learned to request objects, food, and rubs by pressing symbolic keys on a specially designed keyboard. Although ergonomics was not central to this experiment, the construction of the submergible keyboard required a basic analysis of the appropriate means for a large aquatic animal such as a dolphin to interact with a computer interface (Lynn, 2001). In other words, care was taken to ensure the learning was in the task domain and not the interface domain.

3.5 Dog-Human Communication Modes

We have now established that animal-computer interaction must start with an understanding of the animal's basic physiology and psychology, and successful interfaces between animals and computers have taken these considerations into account. Below is a discussion of four sensory interaction modes between dog and human, taking into account the physiology of both creatures. We also evaluate available technology for implementing a remote interaction that supports each mode. From this analysis, we hope to have a clearer understanding of the shape of a remote dog-human interaction.

'Darn these hooves! I hit the wrong switch again! Who designs these instrument panels, raccoons?'

Figure 3-3: © 1986 United Features Syndicate, Inc.

3.5.1 Audio

Canine hearing is at least as good as human hearing, making audio a useful mode of interaction:

> … human hearing usually ranges from a low of 13 – 20 Hz (cycles per second) to a high 16,000 – 20,000 Hz, about seven octaves. Audiologists say that humans hear best between 1,000 and 4,000 Hz … The auditory range of dogs is somewhat greater than that of humans. The region of maximum sensitivity, or best hearing, is 200 – 15,000 Hz. At the low frequencies of 20 – 250 Hz, dogs and humans hear with the same acuity. Above 250 Hz, the dog has a lower intensity threshold to response, and its best sensitivity with lowest intensity is at approximately 8,000 Hz. The upper limit of the canine audible frequency range varies considerably from 26,000 Hz to between 70,0000 and 100,000 Hz (Beaver, 1999).

Not only do dogs possess the apparatus for hearing, they can distinguish between similar sounding human vocalizations. The Ethologist Victor Sarris had three dogs named "Aris," "Paris," and "Harris," and successfully trained the dogs to distinguish between the very similar sounding names without confusion (Coren, 2000).

The frequency response of a typical two-piece PC speaker (no satellite subwoofer) is 70Hz – 20,000 Hz (Sweet, 1999), making these low-cost devices appropriate for rendering sound to both humans and dogs. These speakers clip high frequencies right around the threshold of human hearing, but well below that of dogs. However, they still fall within the boundaries of maximum canine hearing sensitivity. While the relationship between live and rendered sound is probably perceptually different for a dog and a human, there is likely sufficient similarity between human and canine hearing to assume correlation between the ability to interpret sounds coming from a speaker and live sound. The artifacts a loudspeaker introduces are probably not significant.

Audio is attractive because audio transmission is technically mature, can be digitally transmitted using bandwidth reasonably available to the average Internet user, and does not require any special equipment. For these reasons, we expect to make audio a cornerstone of any remote interaction between humans and dogs.

3.5.2 Vision and Video

In contrast to hearing, canine and human vision share less similarity. The most similar feature is binocular vision allowing both dogs and humans to perceive depth and distance. Contrary to what many people think, dogs are not colorblind. Dogs can successfully perform color discrimination. The controversy around dogs and color vision is whether dogs actually use color in their daily lives (Beaver, 1999).

Dogs do not perceive fine details as well as humans, probably because they do not possess the corresponding fine motor control that would utilize such visual acuity. However, dogs are about four to five times more sensitive to low light than humans. Dogs are also most sensitive to moving objects, probably an evolutionary adaptation for hunting. Dogs recognize moving objects at 810 – 900 meters, and stationary ones at 585 meters (Beaver, 1999). Anecdotally, herding dogs are able to correctly recognize a shepherd's hand signals from a great distance, indicating visual acuity on the same order of magnitude as a human. Dogs can also be trained to perform tricks on hand signal cues from their owner.

Although it is clear that appealing to the sense of vision is an important component of the interaction, it is unclear whether current CRT and LCD technology for rendering electronic images means anything to a dog. Modern human-readable displays are highly optimized to specifics of human retinal physiology, and will, in general, fail for non-humans (Fleishman, McClintock et al., 1997). Furthermore, no depth cues are provided with two-dimensional displays – foreground / background images do not move relative to each other with viewer head movement. Unlike loudspeakers that reproduce sound by actually producing sound, electronic displays do not reproduce the object. They rely on visual trickery to reproduce the red, green, and blue components of an object that correspond to our uniquely human perceptual physiology.

It is therefore unlikely a dog can acquire much meaningful information from a computer display device. While dogs can probably distinguish shapes and some colors on a computer screen, they are probably not able to fuse these into meaningful images they can recognize. This fact is clearly stated in "Canine Behavior – A Guide for Veterinarians":

> Most dogs do not watch television, probably for other reasons than the quality of shows available. Although certain sounds tend to draw a dog's attention, the poor form recognition of their vision may make the picture insignificant. It has also been suggested that the scanning rate of the gun in the picture tube, which updates 60 times per second, is too slow to keep a full picture for a dog, which sees only a flicker. In support of this supposition, dogs can discriminate individual flickers of light at rates much higher than humans can. Whereas humans can usually detect 50 – 60 Hz, dogs can detect 80+ Hz (Beaver, 1999).

However, anecdotal testimonials from owners insist their dogs are watching TV with them. There are two explanations for this. First, the dog is simply attending to the sound of the television. Second, the dog is attending to the object of the owner's attention. Dogs have the ability to follow and be directed by a human gaze (Miklosi, Polgardi et al., 2000). Gaze following likely evolved for using non-vocal channels to coordinate pack hunting. Therefore, the center of the owner's gaze is of logical interest to a dog, especially if this behavior is reinforced by human affection and praise. The dog learns that when it looks at the TV, it gets rewarded. In summary, we do not take testimonials of dogs watching TV to be evidence of a dog watching television for the actual content. We have been unable to find any controlled laboratory studies of canine interest in television. However, we must admit the intensity of these television-watching claims makes it difficult to discount them entirely, but in the absence of hard evidence, we rely on the physiological and anatomical studies indicating computer screens are inappropriate devices for dogs.

Beyond the canine physiological shortcomings of modern electronic display devices, design issues surround the inclusion of video. Assuming a display could be modified to utilize components compatible with canine vision, it should not be taken for granted video would play a useful role in remote dog-human interaction.

Videophones for human-human communication have been one of this century's most spectacular technology failures (Egido, 1988). Almost since the invention of the telephone, people have been predicting that the videophone would replace the conventional voice telephone, yet the videophone has made little penetration into the commercial market and virtually no penetration into the home market. Much of this failure is based on economics and technical shortcomings. Videoconferencing systems tend to be very expensive, suffer from poor resolution, poor synchronization between image and sound, and unacceptable amounts of latency. Videoconferencing solutions have also been marketed as replacements for person-to-person interaction, instead of improvements to the conventional phone (Egido, 1988). But despite these logistical shortcomings, one would expect that if humans truly wanted videoconferencing for remote interaction, economical and technically feasible devices would have come to fruition. At best, videoconferencing affords a slightly richer interaction than email or telephone, but does not approach the reported quality of person-to-person interaction (Fish, Kraut et al., 1992). Based on its relative failure between humans, why would one expect it to be useful for dogs?

Isaacs and Tang, however stress video's importance for conveying nonverbal cues: "...relative to audio only, video would also be of use for handling other highly interactive situations when nonverbal cues are most helpful, such as negotiating or creating rapport" (Isaacs and Tang, 1993). Furthermore, many of the shortcomings of video, such as the inability to have side conversations, do not apply to dogs. Dogs do not have side conversations. And because communications with dogs rely greatly on non-verbal cues such as ear and tail position (Coren, 2000), video could be an effective addition to an audio interaction.

In terms of technical feasibility, realtime desktop video interfaces connected to PC computers are fast becoming inexpensive and reliable. With the proliferation of highspeed networks to support video bandwidth, the infrastructure for video transmission is falling into place as well. A typical system will transmit a black & white 160 x 120 pixel image at 60 Hz[6]. Inclusion of realtime video is therefore technically feasible.

From the human perspective, being able to see the dog clearly makes sense. Dog owners acquire information about a dog's emotional state through its body pose, and have the cognitive and optical ability to translate moving images on a 15" screen into a representation of our dog. We have received several reports of people who regularly monitor their dogs over webcams.

From the dog's perspective, the present state of video is less clear (pun intended). Technical considerations aside, dogs are indeed sufficiently visually oriented to benefit from a visual signal. Video fails only when it is considered from the viewpoint of implementation. Furthermore, beyond the perceptual issues, the screen has no depth, and it repositions and resizes images inappropriately as relative orientation changes. For these reasons, video for a dog will probably not be a key component of the interaction, and could even be perceived as a distracting flicker. Therefore, interactions making heavy use of subtle visual signals should probably be avoided.

---

[6] Technical data taken for the 3Com HomeConnect digital webcam: 3Com home connect webcam - http://www.3com.com/products/en_US/detail.jsp?pathtype=purchase&tab=features&sku=003718-00 This camera was top rated by http://www.cnet.com and costs about $125 dollars in 2001.

Although there is no research on developing a canine-appropriate video display, there is work on adding more co-located nuances to video interaction. Systems such as ClearBoard allow collaboration between remote parties using a video system that preserves gaze (Ishii and Kobayashi, 1993). The system approximates two people on either side of a glass screen, both with markers that can be used to write or paint on the board. Given that dogs have the ability to follow human gaze (Miklosi, Polgardi et al., 2000)(and we accept the reverse is true), this ability to communicate focus of attention would clearly be useful to both dog and human. Assuming the technical display hurdles outlined above could be overcome, this would be an useful component of a remote dog-human interaction.

3.5.3 Olfaction

Olfaction is probably the sense for which dogs and humans have the least in common. Dogs have a highly developed sense of smell – reported to be up to 50 times as great as that of humans (Syrotuck, 1972; Pearsall and Verbruggen, 1982). One eighth of a dog's brain is devoted to processing olfactory signals (Syrotuck, 1972). Dogs would likely appreciate the opportunity to receive a "live" smell of their owners while separated. Given that humans can discriminate between the sweaty smell of test subjects watching a happy movie versus a sad move (Chen and Haviland-Jones, 2000), it reasonable to speculate dogs pick up significantly more emotional content from our odor. This is probably even truer when the smell comes from a human a dog knows very well and for whom has had more time to learn associations between smell and mood.

Even though olfaction is not as developed in humans, it is not totally dormant. Babies and mothers can recognize each other with smell alone other shortly after birth (Porter, 1999). Much of the effect of odor on humans is emotional and subconscious. Human subjects exposed to various smells report non-trivial mood changes despite the fact they are just as likely to articulate an aversion to the odor as an affinity (Chen and Haviland-Jones, 1999). Finally, human test subjects have been able to successfully identify their dog through odor alone (Wells and Hepper, 2000). So an interaction with a dog could possible be enhanced by the human's ability to smell their dog.

Unlike color, which (for humans) decomposes nicely into red, green, and blue components, odor has no clean elemental components. English does not even have words to describe platonic odor forms. Linguistic references to odor properties tend to cite instances of odors, e.g. garlic, gasoline, or "new-car" as opposed to categories. The few instances of more universal odor tokens such as "musty" tend to conjure specific images such as an old shoe. This system of odor classification is similar to identifying color only through instances of objects, e.g. "fire-truck color" or "banana color" (Kaye, 1999).

Because odor is such a chemically dependent process, it resists attempts to recast it in an optical or electrical mechanism that can be easily transmitted over distance.

Devices that detect odor are expensive, slow, and have limited generality. A typical sensing product such as the Cyrano Sciences Cyranose 320 detector takes about 10 seconds per sample, and can only compare the sample to a database of previously learned samples. It is intended more for "smell and tell" applications similar in sensitivity to human smell (Cyrano Sciences, 2001). Because of these technical shortcomings a general purpose "odorphone" simply cannot be part of this interaction. Similarly, commercial devices that recreate generalized smells are only beginning to emerge and are not forthcoming with details (Digisents.com, 2001)[7].

Perhaps we can implement a subset of scent? Some research projects have investigated the use of limited scent production to foster a remote interaction. Rob Strong and Bill Gaver developed a system using scent to indicate remote presence of another human (Strong and Gaver, 1996). When one person picks up a picture frame a signal is sent to a remote location. A heating element warms an essential oil, which releases a smell into the environment, letting the other person know they are being thought about. But the scent intensity is difficult to modulate, and is purely one-dimensional, giving the recipient more or less a two-bit signal.

Given a dog's sensitivity to scent, this installation is akin to a remotely controlled tape recorder that is either silent or repeats "I'm thinking of you ... I'm thinking of you ... I'm thinking of you." Such an interaction contains only the barest hint of information. Even though this operates over an acoustic channel, one would hardly equate it with a telephone.

Similarly, Joseph Kaye has developed various scent output devices, including a 2-bit device consisting of peppermint and anise. Even though humans had moderate success in identifying various mixtures, for a dog this would probably be similar to the above implementation, but with two tape-recorded messages instead of one.

Despite the low fidelity of these implementations, they still could nevertheless play a role in a remote dog-human interaction. A small amounts of information is still more information than no information, especially if it appeals to an otherwise under-utilized sense.

3.5.4 Touch

Touch is a fundamental part of the human-animal relationship. Stroking a pet's fur is speculated to be a major component of the relaxational benefits of pet ownership (Beck and Katcher, 1996). When the designers of the Dogz and Catz virtual pet software were faced with creating a realistic interaction with a computer pet, one of the most common questions was "how do you pet a computer animal" (Resner, 2001). The idea of a pet without touch goes against our fundamental notion of pet enjoyment[8].

---

[7] The Digiscent company website (http://www.digiscent.com) says iScent is a speaker-sized device that can emit "thousands" of smells. Little information is given on the variety of available smells.
[8] Fish are a notable exception; Pet fish do not normally exchange touch with humans.

Unfortunately, like remote olfaction, generalized transmission of touch suffers from technical difficulty. But unlike remote olfaction, touch is an important sense to both humans and dogs, and creating a shared sense of touch is a topic of human-computer research (Brave, Ishii et al., 1998).

Human-oriented projects such as inTOUCH and PSyBench (Brave and Dahley, 1997; Brave, Ishii et al., 1998) and LumiTouch (Chang, Koerner et al., 2001) attempt to recreate a sense of touch through interaction with shared objects that maintain a synchronized state. Physical changes to one set of objects are reflected in an identical corresponding physical changes to the peer object. This creates the illusion of directly interacting with the remote person. These devices support non-verbal communication either in isolation of verbal/visual interaction, or as a supplement.

All of these devices, however, rely on the use of manipulatable physical objects to represent something else. In the case of PSyBench, each user has a chessboard like grid on which several tagged pieces are placed. Moving a piece around the board causes the corresponding remote piece to make the identical move, thus maintaining a synchronized state. This interaction is not interesting because users enjoy seeing pieces magically move around a board or derive an intrinsic pleasure from moving pieces, but rather because the movement of a piece represents volition on the part of a remotely located human. If a computer were controlling the pieces, interest would probably not be sustained. Similarly, users engage the pieces on their board because they understand the motions to be similarly interpreted by their remote peer as representing themselves. There is probably not much intrinsic pleasure in simply moving pieces around a board.

Petting is not an intellectual representation of a happy animal, but a primary visceral animal/human response to physical contact. Petting is a very literal interaction between pet and human. Substituting petting by a more abstract physical interaction might give the human some sense of the pet, but it would probably contain very little meaning for the animal. Similarly, expecting an animal to manipulate an abstract object as a way of conveying meaning to its remote owner is well beyond the cognitive capacity of dogs.

Because of animal's limited ability to deal with complex abstraction, touch and petting will not be part of our interaction.

3.6 Asymmetric Interfaces: Creating an Artifact For Human – Non-Human Interaction

Our goal is to construct a device that allows pets and pet owners to have meaningful interactions while geographically separated. In the same way a telephone allows two distant humans to communicate in a meaningful and rewarding manner, we hope to develop a device that performs a similar function for pets and humans.

Figure 3-4a: Face to face human-human conversation

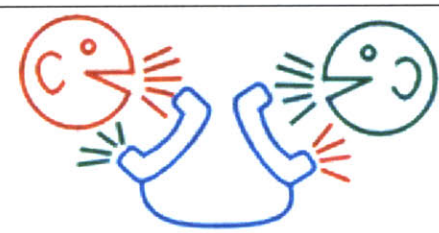


Figure 3-4b: Conversation mediated by technology device drawn in blue.

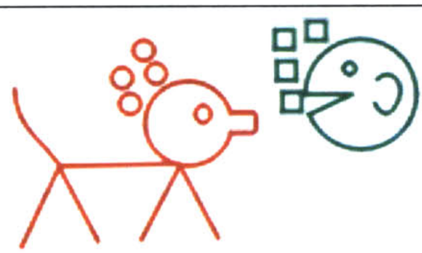


Figure 3-4c: Human-animal interaction. Circles and squares represent the generalized modes of communication each creature produces.

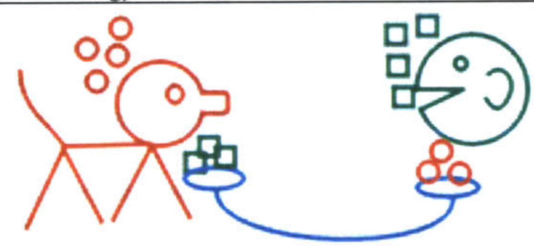


Figure 3-4d: Human-animal interaction mediated by technology device drawn in blue. Developing such a device is the goal of this thesis.

Figure 3-4: Schematic of technology mediated human-human conversation, and how it inspires our search for a similar human-animal interaction that can be mediated by the computer.

The four parts of Figure 3-4 outline our approach. Figure 3-4a shows a human-human interaction void of any technological mediation. Figure 3-4b shows the same interaction, but mediated by a piece of technology (in this case a telephone). The telephone microphone transduces each speaker's voice into an electronic stream and reconstructs a very similar acoustic stream on the other end.

Ideally the telephone is transparent to the interaction – it faithfully transmits all the components of a conversation. In practice, however, not all of the interaction is relayed. In the case of a telephone, one cannot see the other person and cannot interact with shared objects. But a telephone is nevertheless a popular device because enough of a conversation is transmitted that is still feels like social interaction.

For as long as phones have existed, people have probably been putting them up to the ears of their pets and searched for a signs of recognition. Why do pets and humans have a difficult time communicating with a phone? Because there is no real-world precedent for this behavior – pet owners do not have symmetrical back-and-forth verbal dialogs with our pets. Telephones take advantage of a dog's ability to hear, but do not compensate for their inability to speak. Pet owners talk to their pets, and given that dogs have hearing at least as sensitive as human's, it is reasonable to assert they recognize our voice. But they don't respond in kind. They may respond with a howl, tail wag, or purr but (so far) pets do not talk back to us in kind[9]. Symmetrical verbal communication is inconsistent with our task analysis of dog-human interaction.

Figure 3-5: The telephone, which is designed for symmetrical human-human communication, is not appropriate for a dog.

Next time, try Rover@Home!

Photograph ©1998 TCL/FPG International

Figure 3-4c shows how pet owners interact with their pets. The circles and squares represent abstract modes with which each creature expresses itself. For example, the circles could represent a dog wagging its tail or a cat purring, and the squares could represent a scratch behind the ears, the throw of a ball, or verbal praise. The circles and squares are anything pets and humans do to interact with each other. Similar to a human conversation, the human and animal are engaged in an exchange, but at least one party is not using words.

Figure 3-4d shows how we plan to take these inputs and outputs, digitize them, and pipe them over the Internet where they are reconstituted for the remote creature. This creates dual challenges. First is the discovery of some type of sufficiently ritualized interaction that can be electronically transmitted over the Internet. Second is actually building and testing such a device.

[9] Notable exceptions are parrots, with their unique ability to vocalize human speech.

What is notable about this diagram is that the human and animal communicate asymmetrically. What the human gets is different from what the animal gets. Devices for human-human communication tend to be symmetric because most humans have the same physiological and psychological apparatus as other humans. A telephone unencodes information the opposite of how it was encoded to facilitate remote conversation. Both participants in a human-human conversation have the same input and the same output organs. This is not necessarily true for interactions between humans and pets.

3.6.1 Human-Human Asymmetric Interaction

It is tempting to draw parallels between the asymmetrical communication between humans and animals and the unique communication modes that exist between differently abled humans. These differences in ability can be permanent as is the case with deafness, or temporary, as found in a very noisy environment. In both instances, the person has an impaired ability to hear and must communicate through means other than verbal speech.

In the case of deafness, the two basic methods for communicating with a hearing population untrained in sign language are writing or keyboard-based devices such as a TTY, and speechreading (reading lips) for input / talking for output (Also called the "Oral Method") (Communication Technology Lab at Michigan State University, 1995; Zak, 1999).

Use of keyboard-based devices is symmetric. Both parties output information by typing on a keyboard or paper, and both input information by viewing the results. Vision is substituting for hearing, and handwriting or typing is substituted for speaking. There is a one-to-one mapping between the two conversational modes, and symmetry is preserved.

Speechreading / talking is, on the other hand, indeed asymmetric. The deaf person acquires information through visual observation, and emits information by producing sound. This is in contrast to how the hearing person participates in the conversation. But the emphasis of the asymmetrical interface is different. The goal of teaching deaf to speechread and talk is to allow them to better mainstream into a hearing population. A deaf person adept at speechreading and vocalization places little to no obligation on the hearing person to change behavior. Unfortunately, speechreading does not play to the strengths of a deaf person. It requires the speaker to talk slowly, be unobstructed, and be well lit. It is a much less effective communication method for the average deaf person than symmetrical sign language. The burden of communication is almost entirely on the deaf participant. For this reason, speechreading is controversial among many deaf advocates (Nussbaum, 1999).

This is in contrast to Rover@Home. We are not trying to make the dog appear to be a human to an unsuspecting or uneducated human participant. Quite the opposite – we are trying to highlight the dogginess of the dog to the human participant. Our goal is to provide the same quality of co-located dog-human interaction in a remote setting. And we expect both dog and human to share responsibility for any training necessary to make such a remote interaction possible. The communication modes are asymmetric between dog and human, but the deviation from co-located togetherness is in fact symmetric.

Similarly, it is tempting to draw parallels between, for instance, someone talking on a telephone in a car, and someone talking on a telephone in an office or at home. The person in a car is perhaps distracted by traffic, while the person in the office is capable of focusing on the conversation. But the focus of any device to assist the driver centers on bringing the driver to the same level as the person in the office. From simple devices such as hands-free interfaces to more complex devices that automatically mute conversation during difficult driving situations such as merges or sudden stops, the emphasis is on the driver. We know of no commercial devices that help the stationary cognitively unimpaired telephone speaker understand the special needs of the driver.

We therefore find this parallel weak. Without making a value judgment, the goal of technology for people with disabilities is generally to mitigate their differences from mainstream abilities. Steven Hawking's eye movement communicator allows him to talk in a voice that any English speaker will understand. His verbal output may be slower, but no special training is required to interact with him. It is his responsibility to communicate. Our goal with animals is not to bring them up or down to the level of humans, but treat them as unique creatures and symmetrically distribute the responsibility of computer-mediated communication onto both creatures.

3.7 Established Rituals and Task Domains

When designing technology for humans, the interface designer starts by observing what humans do. These are the naturally existing task domains to be recreated in an electronic setting. In the example of the telephone, humans spend a great deal of time having conversations. So it should not be surprising that devices that enable conversational tasks over arbitrarily long distances are popular with humans. Telephones work because they can leverage an existing ritual between people. Beyond the setup procedure of dialing and connecting, every verbally communicative human already knows how to use a telephone – they simply do what they are already doing. A telephone faithfully relies on a cognitive model of interaction and information exchange with another human. No additional training is required. Those unable to use a telephone are likely impaired by more fundamental limitations in the task domain, such as not being cognitively able to sustain conversations. Physical shortcomings to the typical telephone interface can be overcome with a different interface such as a voice or eye-motion dialer, but no interface can compensate for a lack of socially appropriate verbal expression.

Established rituals are the low-tech prototypes for more advanced technological design. They are proven modes of interaction – operational for no other reason than their functionality. Modeling these rituals in a technology-mediated setting makes sense, because much of the conceptual heavy lifting has already been done.

One of the most successful computer interfaces based on established rituals has been the graphical user interface (GUI). It works by taking advantage of the "desktop metaphor"(Apple Computer Inc., 1992) that recreates on the screen the same rituals and interactions users have with a typical office environment. There are filing cabinets, folders, and trash bins. The graphical representation allows direct manipulation -- the user can grasp objects on the desktop and move them to another location, dispose of them, or file them. Desktop interfaces work for a wide cross section of people because the cognitive model required for use of a GUI is similar to the cognitive model necessary for using a real-world desk. The intuitive affordances of a GUI further simplify use. Finally, the rituals for interacting with one's desk are fairly consistent across the desk-using population.

Similarly, a device for remote human-animal interaction should leverage existing interactional rituals. Although it is possible to train the human and/or pet to understand a brand new interaction, our job will be significantly easier and likely more rewarding if the interaction works off an established ritual that requires minimal novelty. Our search, therefore, is to find an animal-human ritual that can be transduced into a digital signal on one end and adequately recreated on the other end. As mentioned earlier, a telephone could be used for the human to communicate with the animal, but is not an appropriate tool for the animal to reciprocate.

In practice, not every aspect of the ritual will survive digitization and reconstitution over an electronic network. The challenge is to pipe through enough of the interaction that it will still be recognizable. It is important that what we build capture the essential feature of the inspiring play pattern. The closer the electronically-mediated interaction resembles the source interaction, the more familiar it will be, and the less retraining will be necessary. Given that dogs have a much harder time generalizing to new context than humans, this is particularly important.

The ideal interaction will be sufficiently ritualized that it can be abstracted away from the physical co-located realm. By being able to take advantage of a relatively fixed set of rules, the job of recreating the essence of the interaction across remote settings becomes much more straightforward. Conversely, the interaction needs to be sufficiently flexible to afford opportunity for improvisation, optimization, and ultimately the expression of a more complex and unique emotional state. An overly ritualized interaction risks tedium and ultimately neglect.

An ideal example of a ritualized and enduring canine interaction is fetch. The rules are simple and universal – human throws ball and dog returns ball to human. Perhaps a game of keep-away is inserted between the dog returning the ball to the proximity of the human, and actually releasing the ball. Additional variants are as numerous as individual dogs. But the essential interaction of fetch is constant. It is safe to assert that any dog owner could play fetch with any dog that shows an interest in the ball or stick. The game might not be as satisfying with a strange dog, but the essential interaction is recognizable to all involved.

Fetch takes advantage of a dog's naturally evolved predator instincts to chase a moving object. Selective breeding has reinforced this behavior in certain breeds of dogs. Fetch is very much an encoded behavior from which humans assume dogs derive an intrinsic pleasure.

Unfortunately, fetch does not easily lend itself to remote interaction. First, there is the problem of safely launching an unattended projectile. But much more serious is the problem of ball return. As previously mentioned, most dogs only return the ball to the proximity of the user, not actually at the user's feet. Without employing complex robot arms controlled with realtime vision analysis, an attainable computerized fetch relies on the dog's ability to successfully place the ball in a designated receptacle. As mentioned before, this step is not part of a typical dog's fetch repertoire. Only in remote fetch, does this previously trivial step become critical. This is not to say fetch is out of the question. With training, the shortcomings of fetch can be overcome.

3.8 Low-Tech Prototype: Clicker Training

With these caveats and design goals in mind, we set out to find an interactional ritual that dogs and humans find inherently rewarding, and could be implemented with currently available technology. As mentioned before, fetch is a great ritual, but technically challenging.

The ideal interaction will use audio for both human and dog, video for the human, olfaction for the dog, and as much touch as technology will allow.

3.8.1 Clicker Training Background

Clicker training is an established animal training technique that associates the sound of a toy clicker with a food reward. Clicker training has its roots in the work by B.F. Skinner on operant conditioning in the 1940s(Burch and Bailey, 1999). Two of Skinner's students, Keller and Marian Breland, were the first actively to commercialize Skinner's academic work and created a successful business "manufacturing" clicker-trained animals for zoos, parks, and fairs. The word "manufacturing" is used because of the engineering cookbook approach the Brelands took in their approach to clicker training (Breland and Breland, 1951; Breland and Breland, 1961). They felt they had discovered truths about animal behavior that could be used systematically to shape any behavior an animal is physically capable of performing. Included in their product line were chickens that play tic-tac-toe and ducks playing basketball (Burch and Bailey, 1999).

It was not until 1984 when a retired dolphin trainer, Karen Pryor, wrote "Don't Shoot The Dog" (Pryor, 1984) that clicker training became an option for amateur dog owners (Wilkes, 1995). "Don't Shoot The Dog" was the first non-academic book to apply the operant conditioning work of Skinner in a casual and populist style. Its relevance was not just to dogs and other animals, but to humans as well.

In 1992 dog trainer Gary Wilkes started using clicker training as a method to control and shape dog behavior (Wilkes, 1995). Dog trainers became excited about clicker training because it offered an effective method of dog training that did not center on punishment or other aversive stimulus. Beyond ethical considerations, they also found clicker training to be a highly effective means for controlling and shaping a dog's behavior. This thesis revolves very tightly around the work Wilkes has done with clicker training.

3.8.2 Clicker Training Process

Clicker training unfolds in three basic steps (Wilkes, 1995; Spector, 1999). The first step is to associate the sound of a toy clicker with a food reward. At first dogs are likely to be startled by the sound of a clicker, but even the most skittish of dogs quickly learns that the sound of a click is an accurate predictor of a food treat. According to traditional Pavlovian conditioning, the click becomes just as good as a food reward, and elicits the same expectant behavior response of salivation and begging as does actual food (Gray, 1994). A dog conditioned to the clicker should expectantly look for a treat upon hearing the click sound. Dogs new to clicker training will often ignore or be startled by the click. It is important to note there is nothing special about the click. Pavlov used a bell, and dolphin trainers use a whistle (Pryor, 1984; Ramirez, 1999). Clicks are good for dog training because they can be heard over long distances, are easy to consistently administer by anyone with a clicker, have a unique and unmistakable sound, are not modified by emotional infection, and are very short duration (Spector, 1999).

The reward should be anything the dog enjoys, but is typically a food treat. As long as the dog understands the click to be an accurate predictor of some type of reward, the clicker is considered to be "charged." This means the clicker sound has meaning to the dog beyond a staccato ambient noise. The motivational significance of the clicker can be increased if working with a hungry dog, and for this reason, many clicker trainers give a portion of the dog's evening meal as part of a clicker-trained dinner appetizer.

The second step is to lure a behavior such as "sit". This is done iteratively, by rewarding successive approximations towards a full sitting behavior. As soon as the dog makes the slightest sitting motion, click and treat. After a few iterations, the dog will learn that lowering his hind-quarters earns a reward and will eagerly repeat this behavior. By slowly raising the threshold for a click, the dog learns he must come lower and lower to the ground until the dog is eventually fully seated.

In practice, simply waiting for a dog to initiate a sit can take way too long – a dog may lose interest well before he offers enough sits for a reward pattern to emerge.

Clicker trainers use a process of luring to hasten results. Holding a piece of food in front of the dog and slowly moving it over and behind the dog's head will cause the dog to back up and lower its hind-quarters. At this point the trainer clicks. By luring the dog into position, the trainer greatly increases the opportunity to reward the dog. As the behavior becomes more established, the lure can be faded.

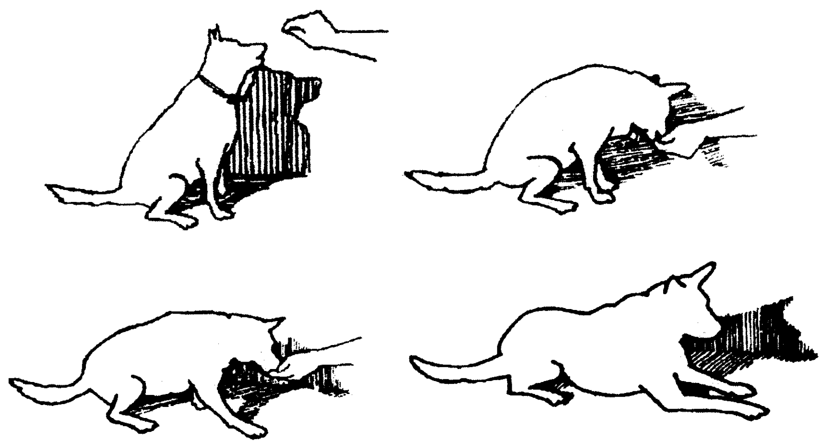


Figure 3-6: Using food to lure a dog from sitting to lying down. Iterative clicking at intermediate stages between sitting and down will keep the dog interested and hasten the learning. While the trainer uses a lure, the trainer never actually touches the dog or forces the dog into position.

Drawing ©1995 Gary Wilkes , Click&Treat Training Kit, Version 2.0.

Another common luring technique is a target stick. Dogs that are trained to follow the tip of the target stick can be lured into any variety of configurations, which can then be rewarded. How to teach a dog to follow a target stick? Clicker training, of course.

The third and final step in clicker training is to add a discriminative stimulus -- extinguish the behavior in a clicker-training session in the absence of a verbal cue. At the end of step two, one has a dog that constantly sits in anticipation of a treat. But the goal of the training is a dog that sits only when told to sit, and does not sit during a session when not told to sit.

Even though the trainer has probably been saying "sit" while luring the dog into a sitting position, in this step the verbal command becomes significant. The dog learns the command "sit" is not background noise, but an opportunity to exhibit a behavior that will elicit a reward. This creates a behavior on a cue, and extinguishes the behavior in the absence of the cue.

As with the clicking sound, there is nothing special about the utterance "sit." Any uniquely perceivable external stimulus such as a hand signal, an odor, or sound may be employed as a signal. As long as the dog can recognize the signal, it can be used to elicit the behavior. Often the trainer thinks the dog is attending to the verbal cue, when in fact, the dog is responding to a hand signal, facial tic, or some other unconscious gesture.

Proponents of clicker training assert that it not only creates obedience, but is also a fun and exciting activity for dogs. Dogs come to understand clicker-training sessions as an opportunity for food reinforcements and social interaction with owner. A happy and excited dog is essential for clicker-training success. Because the dogs must discover successful behaviors on their own, dogs unwilling to explore will have a very difficult time creating opportunities for rewards. Dogs that are punished or subject to aversive stimulus are more likely to display stereotypical behavior such as crouching, cringing, and hiding.

A key feature of clicker training is the lack of physical contact between dog and owner. Traditional dog training techniques that shove the dog's hind to the ground while saying "sit" teach the dog the word "sit" means "I'm going to shove your hind to the ground now" (Pryor, 1984). With clicker training, it is essential the dog discover the behavior on its own. Luring only gives the dog clues; the trainer still relies on the dog's voluntary cooperation to discover the desired behavior. No amount of luring will ever get a frightened dog to cooperate. A dog that becomes excited during a clicker training session is going to be making a large effort to discover the behavior that causes its owner to click and treat. Often trainers have no clear goal in mind – they simply watch their dog and click when the dog does something new and interesting.

3.8.3 Paradigm versus Implementation: Varieties of Clicker Training

As clicker training has grown in popularity, there have been inevitable splits of doctrine. Most significant is the justifiable application of punishment or aversive stimulus. Some clicker trainers believe aversive stimulus such as striking, yanking, shocking, or scaring the dog is never appropriate, and with proper training, such as teaching an incompatible behavior, all unwanted behavior can be extinguished without punishment (Pryor, 1984). Other trainers are more liberal in their application of punishment, stating when used appropriately it is a humane means to control an animal that would otherwise become unsuitable for domestic life, and ultimately be destroyed (Wilkes, 2001). It is unclear what percentage of the 2 to 18 million dogs killed in the United States in animal shelters each year (Beck and Katcher, 1996) are due to behavioral problems, but there is clearly anecdotal evidence to support uncontrollability as a factor in the decision to bring a dog to an animal shelter.

Clicker trainers also disagree about the removal or fading of the clicker sound (Wilkes, 1995; Ramirez, 1999; Spector, 1999). Some feel the clicker sound must always be present, while others adopt a "crossover" approach, transitioning into more

traditional training techniques after using clicker training to shape and mature the trick.
Further ideological differences center around the use of continuous versus variable reinforcement, and the associated practice of jackpotting – giving the dog an occasional huge reward -- "Extras for Excellence"(Wilkes, 1995).

These debates focus on the implementation of clicker-training and are a healthy component of any developing scientific methodology (Kuhn, 1962). Even though trainers differ about how to implement clicker training, all are working within the clicker-training paradigm. All agree that clicker training and positive reinforcement is an optimal means for communicating intent to a dog and therefore an ideal method for shaping new behaviors.

This thesis and associated artifacts fully accept the clicker-training paradigm and explore the use of clicker training in settings where dog and owner are not co-located. But we do not add to or comment on the debates within the clicker training community. We are building devices that allow the clicker training paradigm to be performed remotely. We do nothing to insist on any particular implementation. We simply provide the tools and methodology for remote clicker training. It is up to each trainer to decide what to say, when to click, and when to treat.

Although we have not created any aversion devices, nothing in either the technical implementation or design methodology precludes the use of aversive stimuli, such as citronella spray, mild electric shock, or startling noise. But building such devices would be an implicit endorsement of their use, whereas omission allows us to remain neutral.

It is our intuition that canine behavioral problems should not be dealt with remotely. Behavioral shortcomings are a delicate issue best resolved in a co-located situation where both dog and owner have full benefit of all nuances of facial and body expression, odor, sound, and especially touch. Rover@Home has is not intended to treat, diagnose, or remedy any condition that is best addressed by a qualified professional canine behavioral consultant.

3.9 Clicker Training as Dog-Human Communication

Clicker-training creates a communication pathway between human and dog. The click becomes a precise means for the owner to communicate intentionality to their clicker-trained dog. Properly administered, the clicker sound gives the dog the following two pieces of unambiguous information:

- **Event Marker:** What exactly the dog did to earn the reward. The clicker marks the exact point in time of optimal behavior rendition.

- **Behavioral Transition:** The above behavioral feedback also signals the trick is over, come get your reward. Some clicker trainers use the click to mean:

"getting close, keep going," but this message conflicts with the click as a predictor of reward (Spector, 1999). How can a dog both continue a trick and come and get a reward? As a conditioned reinforcer, the click must be temporally associated with a treat to retain meaning.

Morgan Spector succinctly states the communicative ability of the clicker in his book on clicker training:

> " … [clicker-training] rests on a training 'bargain' between you and the dog. Simply stated, you say to the dog, 'You give me what I want, and I'll give you what you want.' The training process then becomes a dialog between you both. The dog looks to you for direction (information about what you are asking for) and confirmation (that the dog gave it to you). It is your job to be clear, consistent and far.
>
> "The click sound made by a hand-held clicker is the primary way that you 'talk' to the dog in the training process. It works because you consistently deliver what the click promises, so the dog becomes comfortable with its unambiguous message" (Spector, 1999).

Much dog-human interaction, especially with training, revolves around praise. But praise is sloppy information. Praise builds esteem and confidence, but it does not provide the dog with knowledge about what specifically it did to earn a reward. Praise can be the unconditioned stimulus associated with clicker training, but it is hard to make it information. Clickers are ideal at conveying precise information that is the foundation of a dialog.

A popular joke with clicker trainers says that the human thinks it is training the dog, but the dog thinks it is training the human. The dog asks: "what do I need to do to get this silly human to give me a treat."

3.10 Clicker Training as Computer Interface

Several features of clicker training make it ideal for use in a setting where dog and owner are not co-located. The modes of interaction between dog and owner are highly ritualized, and are amenable to transmission over a digital medium. There is a large installed base of clicker-trained dogs that are already prepared for Rover@Home.

Most importantly, clicker training does not fundamentally involve contact between dog and human. In fact, versions of clicker training which most closely adhere to the original works of B.F. Skinner highlight the importance of not manipulating the dog. Only if the animal is allowed to discover the behavior on its own will it learn (Burch and Bailey, 1999). Physically moving the dog into position is viewed as "cheating" and contrary to the learning process. While most clicker trainers are more relaxed about this dictum and actively pet their dog as a reward, being able to constantly touch a dog is not a fundamental component of clicker-training.

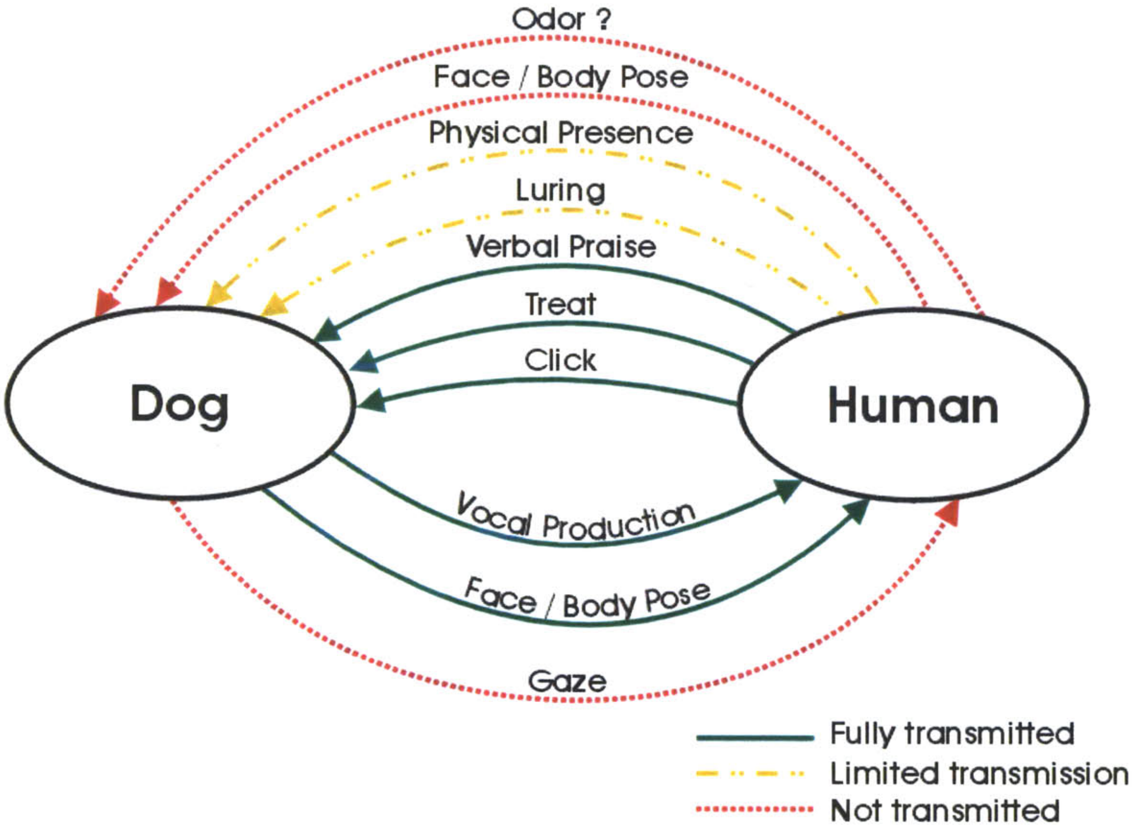


Figure 3-7: Communication modes between dog and human using clicker training. The solid lines (green) are reproduced with Rover@Home, while the dotted lines are lost when the computer mediates the interaction (red). The dot-dash lines (yellow) are only partially reproduced.

3.10.1 Human to Dog Communication

The owner's primary means of communicating with the dog is through the precise delivery of a clicking sound. This is the signal that tells the dog its owner is pleased and the dog is about to receive a food reward. Delivery of a food treat is another mode the owner uses to communicate with his dog. Verbal praise along an audio channel is a third mode. All three of these modes are easy to control over a remote network.

The owner also communicates with the dog through luring, either with their hand or with a target stick. Unlike remote delivery of clicks, food, and verbal praise, arbitrary motions in three-dimensional space are not easy to implement. But while luring to any object along any path is difficult, subsets of luring are possible. If the goal of luring is to give the dog a clue as to the desired behavior, then flashing lights and sudden sounds are sufficient for many task domains.

Our hypothesis is the absence of fully generalized luring will impair the dog's ability to learn completely novel tricks, but will have a diminishing effect as the desired behaviors become more familiar. In the absence of novelty, luring becomes almost unnecessary. A dog that knows how to sit should be able to sit without any luring at all.

The owner also communicates through display of visual cues. Many owners believe their dogs are responding to verbal cues, when, in fact, they are responding to subtle hand signals or body cues (Ramirez, 1999; Spector, 1999). If the intent of clicker training is to prepare a dog for remote interaction, it becomes important to avoid superfluous movement so that the dog truly learns the importance of the verbal cue.

Analysis of clicker-training videos shows that during clicker training sessions, dogs are primarily focused on their activity and not looking at the trainer (Pryor; Wilkes, 1995). Dogs look at their trainers between tricks, when expecting rewards, and when frustrated. But in general, while discovering new behaviors, dogs seem primarily focused on their own activity.

Olfaction is also a likely component of the interaction between dog and owner. As mentioned earlier, given the sensitivity of a dog's nose, dogs very well may be able to gather information about the owner's emotional state through the realtime emission of pheromones or other body secretions. Unfortunately detection and reconstitution of odor is not technically feasible.

Finally, the owner communicates through physical presence. Even if the owner is silent and still, the owner's existence in a shared space can be an indicator that a clicker training session is in progress. This presence alerts the dog to be expectant of verbal commands, clicks, and treats, and not to wander off or be distracted.

3.10.2 Dog to Human Communication

When dog-human communication is studied, a similarly straightforward interaction is found. The dog communicates by performing the desired behavior. The owner typically evaluates acceptability by visually inspecting the behavior and rewards as appropriate. Therefore, if the trick involves a body pose, a video channel from dog to human is essential, and audio actually becomes superfluous in terms of the owner's ability to competently evaluate successful trick completion.

However, if the trick exclusively involves vocal production on the part of the dog, audio becomes fundamental and video becomes secondary. If an owner is content only to train vocally oriented behaviors such as "speak" or "whimper", higher-bandwidth video becomes superfluous and only audio is necessary.

In practice, though, owners are likely to want both audio and visual input. Video without sound feels eerie, and audio without video interferes with the owner's ability to create a sense of interaction with the dog. We believe owners want to see their dogs, and hear the ambient sounds dogs make – even if these cues are not essential to the core interaction.

4 Design and Issues of Sample Implementation

4.1 Implementation

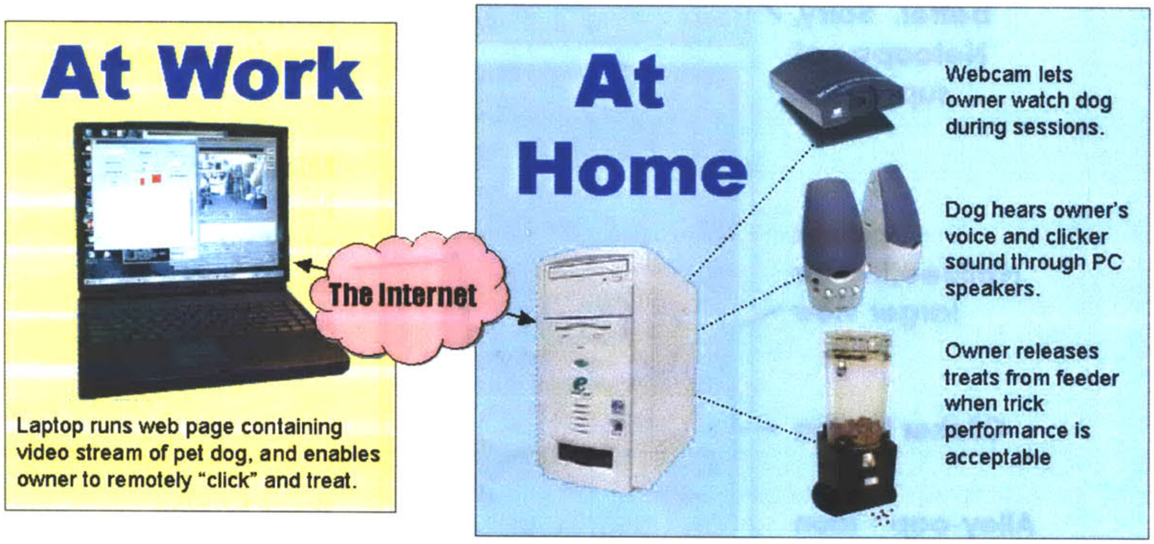


Figure 4-1: Minimal Rover@Home setup. More advanced version include instrumented props which aid the remote interaction and help make it more similar to co-located clicker training.

The minimal Rover@Home setup is outlined in Figure 4-1. At the user's home is an Internet-connected computer with a webcam, PC speakers, and a custom-made food dispenser. The food dispenser is the only custom hardware in this basic setup. All other components are off the shelf and can be obtained from a variety of vendors. Microsoft NetMeeting is used for the audio/video link between the home and remote computers. The dog's computer runs custom software that serves a web page containing the controls for operating the Rover@Home equipment.

The remote or traveling computer has no special hardware or software requirements. The computer must have Internet Explorer 5.0 and Microsoft Net Meeting (which is free to download and installed by default with Windows 2000). Therefore, any recently updated Windows computer should be able to act as a Rover@Home client. This ease of use was an engineering cornerstone, included to make Rover@Home as convenient and ubiquitous to use as possible. We want users to be able to interact with their dogs from a variety of locations, including work, client locations, friend's homes, and Internet cafes.

Care was taken to keep any dangerous voltages well away from the dog. All cords attaching Rover@Home peripherals to the computer are standard 4-wire RJ-12 phone cables, which are harmless to the dog if chewed and readily replaceable from a wide variety of sources such as Radio Shack.

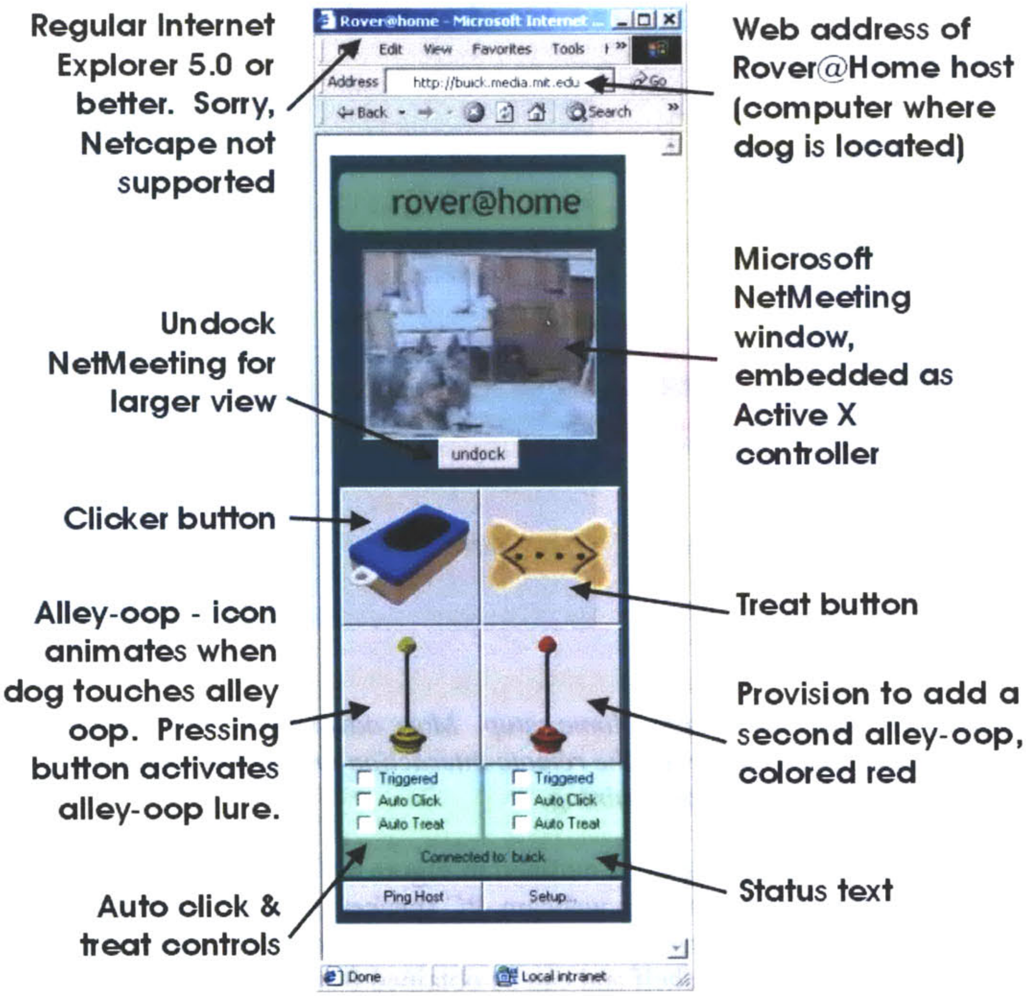


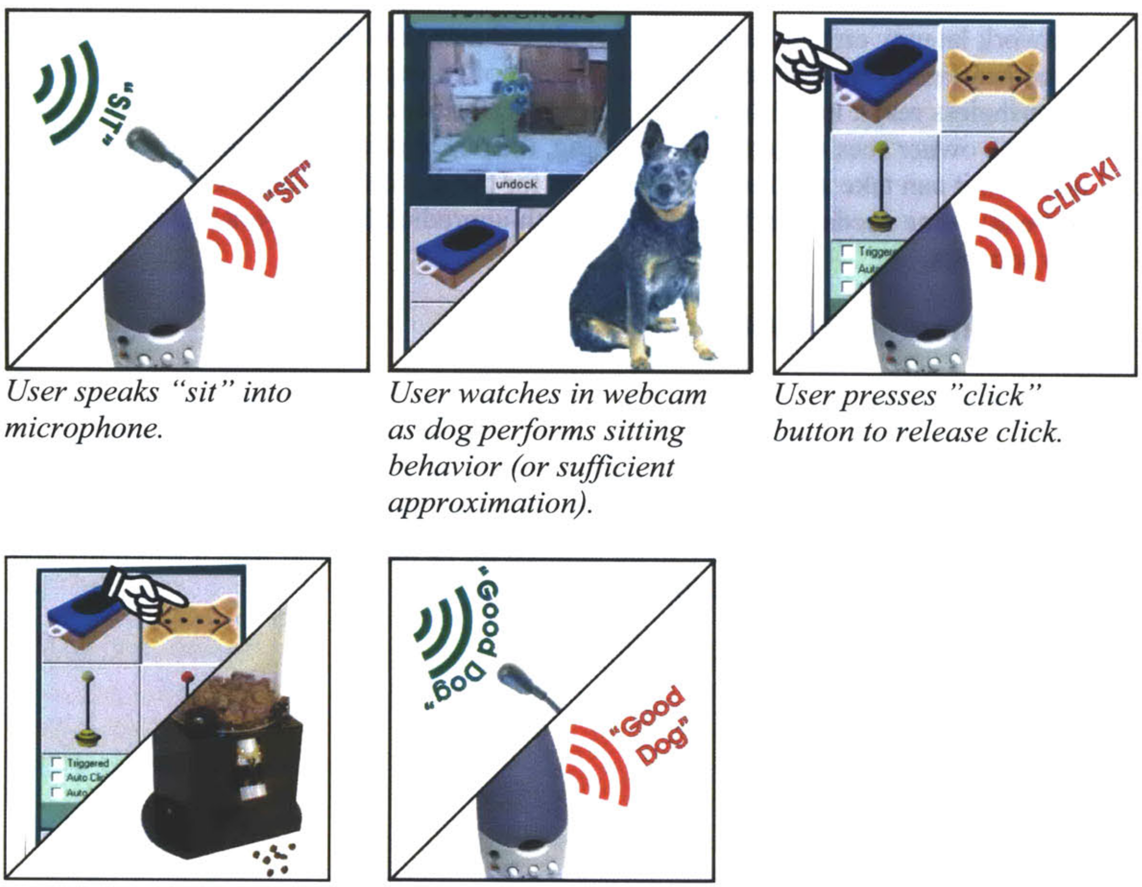
Figure 4-2: The Rover@Home web page used by the remote user. Any computer with Internet Explorer 5.0 or greater and Microsoft NetMeeting (free to download, installed by default on Windows 2000) can communicate with the Rover@Home computer. Thanks to Scott Eaton for the graphic design of this interface.

Figure 4-2 shows the Rover@Home web page the user employs to interact with the user's dog. The user types in the URL of the home computer where their dog resides. This can be either a fixed IP address or a name that will be resolved into an IP address. The custom software on the home computer serves the Rover@Home webpage and establishes the Microsoft NetMeeting link.

The user can click and treat, just as in the co-located interaction. A video window allows the owner to view the dog. All controls provide visual feedback that the action has been taken. Furthermore, the treat icon displays a warning if the treat dispenser is empty or was otherwise unable to release a treat. A malfunctioning treat dispenser could easily jeopardize clicker-training.

Interactions with Rover@Home proceed very much like co-located clicker training sessions. Just as in co-located clicker training, the user initiates sessions with a few clicks and treats to "charge" the clicker. The user watches the dog in the webcam window, and once the dog is properly attending to the clicks, the training session can begin. The user speaks a command into the microphone, where it is played to the dog. If the dog's behavior is sufficient, the user clicks and treats. For instance, the user would say: "sit," and if the dog sits, the dog gets a click and a treat.

The red and yellow "alley-oop" buttons and checkboxes underneath are used in conjunction with more advanced features, explained below. Alley-oops are props that dogs are often trained to "touch," which means they touch the tip of their nose to the tip of the alley-oop.



User speaks "sit" into microphone.

User watches in webcam as dog performs sitting behavior (or sufficient approximation).

User presses "click" button to release click.

User presses "treat" button to release treat.

User praises dog for job well done.

Figure 4-3: Basic Rover@Home interaction.

Multiple users can run concurrent Rover@Home sessions, but only the first user can get video. Furthermore, if one user treats, the other user receives symbolic information about that action. The states of all users control panels are kept synchronized by the server software.

4.2 Implementation Issue: Potentially long latency

One of the most difficult aspects of clicker training, especially for new trainers, is the careful synchronization of the click with the desired behavior. Missing the critical moment by even a second is enough to confuse the dog as to the precise behavior that elicited the click and associated reward. It is very important to click at precisely the moment of optimal behavior. Many novice trainers intellectualize and even discuss the dog's behavior before issuing a reward, losing the moment and rewarding the incorrect behavior.

Network latency on the Internet (or any shared network) only compounds this delay. Even near-instantaneous reactions on the part of an experienced trainer can nevertheless result in a delay between trick performance and click delivery. By the time the owner sees what the dog is doing, the dog is already seconds past the trick, and then it can take another second or two for the click and food to reach the dog. Rover@Home needs to effectively deal with unpredictable network latency up to a few seconds.

4.2.1 Theoretical Solution

What if the dog's compliance could be determined by methods other than remote human observation? What if the computer could be used to administer perfectly timed clicks and treats? Instrumenting a dog is awkward and difficult, but monitoring interaction with instrumented toys is quite feasible. Instead of visually observing obedience, the dog's behavior can be inferred by its interaction with a wired toy and perfectly timed rewards delivered appropriately.

Figure 4-4a: Alley Oop detects tilt
Alley Oop © 2001 Gary Wilkes
http://www.clickandtreat.com

Figure 4-4b: Squeaky Toy detects squeeze

Gary Wilkes uses an "alley-oop" (Figure 4-4a) device as a prop in his training regime. Alley-Oops are the first use of non-edible target in animal training (Wilkes, 2001). Dogs are trained to "touch" the tip of the alley-oop, or "plunk" the base with their paw. Other variations are certainly possible.

An alley-oop could be wired in several ways. Most basic is a tilt sensor that determines when the dog is actually moving the alley-oop. This is a binary "tilt" / "no-tilt" signal that is unable to discriminate how close the dog is to the alley-oop or how much it is displaced. More sophisticated sensing includes a proximity detector at the tip that can give information about how close the dog is to the alley-oop. This incremental information is useful for shaping a "touch" behavior where the owner sets a threshold and the computer automatically rewards the dog simply for approaching the alley-oop.

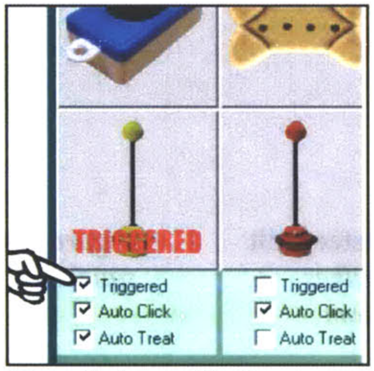
Similar to the Alley-oops, squeeze toys fitted with air pressure sensors can transmit whether or not they are being squeezed. Dogs can be trained to pick up a specific squeeze toys in response to verbal commands. When the dog grips the squeeze toy in its mouth, pressure sensors will relay this information to the owner. An example of squeeze-toy interaction would be "get panda" or "squeeze panda."

An alternative to instrumented toys is to instrument the dog's environment. A fixed camera with a vantage of the dog's environment could feed into simple image analysis software to determine the dog's location through motion energy. When the computer determines the dog's interaction is above an owner-set threshold, clicks and treats are delivered.

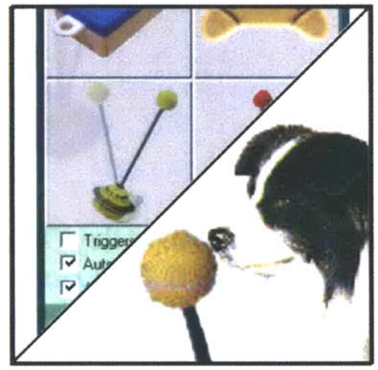
The use of an instrumented environment or toy places necessary restrictions on the behaviors an owner can train. For instance, with an instrumented environment, tricks such as "run in circles" or "go to corner" are easy problems for vision analysis, while more standard dog tricks such as "sit," "paw," and "down" would be significantly more difficult for today's vision analysis software, especially if there is no guarantee of the dog's orientation and position in the room.

4.2.2 Implemented Solution

User speaks "touch" into microphone



User "triggers" alley-oop. It will automatically auto click and auto treat next time alley-oop is displaced.



When dog touches alley-oop, the user's alley-oop icon animates...

(Note "triggered" UI is automatically cleared)

...and the computer automatically clicks...

... and treats.

User praises dog for job well done, just like before. For the dog, this is simply a well-timed interaction.

Figure 4-5: Solution to latency and bandwidth. Computer automatically clicks and treats when dog touches alley-oop. User still gives the dog verbal praise, decides which behavior to work on, and how long the session lasts.

For this thesis, only the alley-oop with the basic tilt sensor, and squeeze toy panda has been implemented. In Figure 4-5 above, the three checkboxes under each alley-oop control the automatic deliver of clicks and treats. When the "triggered" checkbox is selected, the next time the dog displaces the alley-oop, the computer releases a click and a treat. The "triggered" checkbox is then automatically cleared to prevent the dog from receiving additional treats for touching the alley-oop. Remember, the goal is for the dog to only be rewarded if the behavior is preceded by a verbal or other unique command. Simply touching the alley-oop is not good enough; the dog needs to touch the ally-oop because he was told "touch." The user needs to re-check the "triggered" checkbox for each instance of "touch."

If the owner so desires, he can automate either clicks or treats and perform the other part manually. The checkboxes "Auto Click" and "Auto Treat" control what the computer does when the alley-oop is displaced.

The Rover@Home control panel also includes a "Ping Host" button that times the round-trip time of a packet of information. This gives a good gauge of expected network latency. In practice, we have found latency to be anywhere from almost zero to about one second.

4.2.3 Total Automation

While the computer controls the precise delivery of clicks and treats, the owner still retains overall control of the training session. The owner determines when to focus on a difficult trick, when to move on to a new trick, and when to end the session. The owner draws on his extensive knowledge of his dog's particular temperament and learning style to customize each training session. Additionally, dogs are extremely sensitive to emotional inflection in their owner's voice. Because verbal praise is live for each trick, it is highly tailored for that exact situation.

A fully automated interactive clicker training system is definitely an interesting research direction, but it will not be pursued in this thesis. The challenges of a synthetic trainer than can learn the nuances of the dog's moods and deliver appropriate motivational cues are tremendous, and deserve an independent research effort. The goal of this thesis is to enrich the experience for both the owner and the dog. We do hope, however, this work assists future efforts to develop a more automated system.

4.3 Implementation Issue: Large Bandwidth Requirements

The use of video requires a relatively large amount of bandwidth. For home users, this means at least a clean 56k, and most likely a cable or DSL connection. Slower dialup connections simply do not support the video resolution necessary to support interactive dog training. Most current generation Internet-enabled handheld devices are well below the minimum bandwidth cutoff, with maximum data rate typically not higher than 19.2k baud[10].

The instrumented toys developed above provide an elegant solution to augment or replace a video stream. By transmitting symbolic information about the instrumented

[10] Figures taken from wireless Palm VII, https://store.palm.com/Catalog/ProductDetailsChild.asp?ProductNr=3197

state of the toys, the owner can infer what the dog is doing. In the implementation of the alley-oop, the alley-oop icon animates back and forth when the dog touches it. Therefore, owners do not need to actually see their dog touching the alley-oop, they can reliably determine this from their Rover@Home control panel.

Taken to its extreme, these instrumentations could potentially remove the need for video. This would allow Rover@Home over lower-bandwidth low-resolution screen devices such as wireless Palms or web-enabled cellphones. An area of future research would be to investigate the degree to which video could be removed and Rover@Home still provide a meaningful interaction.

Figure 4-6: In the future, this man will use his web-enabled cellphone to interact with his dog over a 19.2 k baud connection.

4.4 Implementation Enhancement: Luring

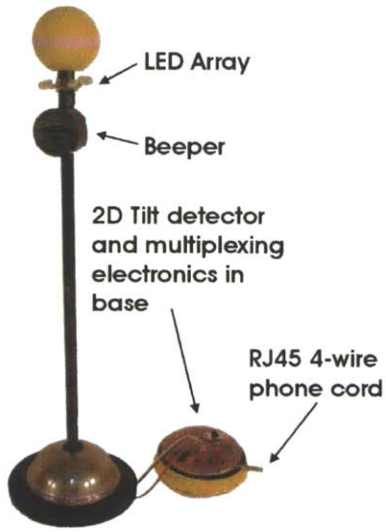


Figure 4-7: Alley oop instrumented to overcome issues of latency, badwidth, and luring. LEDs and beeper can be remotely activated to attract dog's attention, tilt sensors in base detect motion.

Alley Oop © 2001 Gary Wilkes http://www.clickandtreat.com

As mentioned before, arbitrary luring in three-dimensional space is beyond the scope of this thesis. But luring to one of the alley-oop or squeeze toys described above is quite possible. An alley-oop or squeeze toy fitted with lights and / or noisemaker can be remotely actuated, just like the treat dispenser. This will attract the dog's attention, which is a behavior that can be clicked and rewarded. Successive iterations of lure – click – treat can teach the dog to look at the desired object. It is hoped this interaction can be shaped from "look-at" to "touch" by requiring the dog to take steps towards the object in order to receive a reward. Natural curiosity, combined with a clicker-trained dog's desire to explore new behaviors might be sufficient to entice to dog to approach the target object.

Furthermore, it is hoped that after a dog has learned to touch a few different props by following flashing lights and pulsing sounds, it will be able to generalize that flashings lights and sounds means "reward for approach." This is an interaction that can be strengthened in a co-located setting.

Luring, however, is somewhat vulnerable to the latency and bandwidth requirements outlined above. Because at first the interaction with the prop is nothing more than a gaze, machine evaluation is difficult. And because this gaze is likely very short duration, precise clicking is essential. Alley-oops with proximity sensors can detect change is distance from the prop, but they cannot determine gaze or stance.

Instrumenting the props to detect eye gaze through optical evaluation of retinal pigments or corneal reflection is technically possible, but outside the scope of this thesis[11]. This sensing would allow the computer to determine glancing at the object, and automatically reward this first step towards "touch." Successive approximations

---

[11] The 'red-eye: phenomena associated with flash photography is the basis of an effective gaze detection module.

could be evaluated by combining eye gaze detection with physical location in the room, either through proximity detectors on the toy, or a fixed camera monitoring the entire room.

For this thesis, an alley oop with a tilt sensor, LED array and beeper was constructed (Figure 4-7). Also constructed was a wireless squeeze toy that could be remotely triggered to emit a panda-bear like sound.

4.5 Implementation Issue: Poor Timing

Most new clicker-trainers require time to learn the subtle art of clicker training. They need practice to deliver well timed clicks, as well as learning what motivates their dog, and how long their dog is likely to remain interested and focused in clicker training. Rover@Home is simply a computerized interface to clicker-training. It does not teach clicker-training, but allows a remotely located clicker-trainer to perform a similar interaction at a distance.

Computerized control of instrumented toys alleviates some of the burden of accurate timing, but the trainer still has to direct the overall route of the clicker-training session. Trainers who cannot perform this in co-located settings will likely have similarly disappointing results with Rover@Home. Rover@Home competency still resides in the original task domain of clicker-training.

5 Results

5.1 Modes of Evaluation

5.1.1 Human Reporting of Dog Interaction

Subjective words such as "enrich" and "meaningful" are extremely difficult to apply to animals. How can humans determine if their animals do indeed "enjoy" or are benefited from the devices we have built? Perhaps the easiest way to determine the value of Rover@Home for dogs is to ask dog owners. Because we are working with pets, and not wild animals, we can expect the human to be a guardian who not only feeds and cares for the animal, but also has the ability to articulate what the human sees as their pet's inner mental state.

Clinton Sanders, in his sociological book on the relationship between humans and dogs, eloquently states: "Because the animal is 'mute', caretakers often find themselves in the situations in which they must 'speak for' their nonhuman companions. In so doing, they make use of a rich body of knowledge derived from an intimate understanding of the animal-other built up in the course of day-to-day experience. Dog owners commonly give voice to what they perceive to be their animal's mental, emotional, and physical experiences." He does not seem troubled that his insights into dogs come from the mouths of their owners and takes these evaluations at face value (Sanders, 1999).

Because our trials have been mostly exploratory, owner reporting is our primary method of evaluation. We assume owners advocate in the best interest of their pets, and would not praise a device that causes their dog distress or harm.

5.1.2 Professional Evaluation

We also allowed professional dog trainers to use Rover@Home and record their opinions. These experts have the experience to provide insights into the unknown behavioral effects of devices such as Rover@Home. Dog training experts also can articulate technological and design improvements.

5.1.3 Correspondence With Co-Located Activity

We can completely avoid thorny issues such as animal enjoyment, and simply show a behavioral correspondence between co-located and remote interactions. Instead of evaluating whether or not Rover@Home is a worthwhile activity, we simply show that it creates an interaction similar to clicker training, upon which it was based. We leave it as a leap of faith on the part of the owner, that their dog enjoys clicker training. We make no claims as to whether or not the dog "enjoys" being clicker trained, and avoid the debate of animal consciousness. We simply show the degree to which Rover@Home reproduces clicker training in a remote setting.

To some degree, the methodology section of this thesis has been concerned with this endeavor. We have taken pains to show the modes that are reproduced in Rover@Home, and the modes that are lost. We then attempt to analyze the effect of these omitted modes on the overall interaction. None of this analysis rests on any assumption of enjoyment or pleasure.

5.1.4 Physiological Metrics

The most quantitative evaluations are physiological. Measuring data such as gaze, ear alertness, blood pressure, wags per minute, or stance can be interpreted to represent an emotional state. These are also useful for determining correspondence between co-located clicker training and Rover@Home.

We do not explicitly record physiological data, but expect observable canine features are qualitatively incorporated into professional and owner evaluations. An evaluation of fear, for example, is in large part based on observations of body pose and respiration rate.

5.1.5 Elimination of behavior problems

Dogs subjected to social deprivation exhibit the same emotional problems as humans in similarly stressful situations (Beaver, 1999). Acting out behaviors such as chewing or urination become more commonplace in these situations. Many owners associate these incidents with coming home late or weekend trips away from their dog.

Demonstrating that Rover@Home eliminates these behaviors would reinforce our assertion that Rover@Home successfully creates a rewarding social interaction. If an owner could continue to work late, but measurably mitigate behavioral problems with a few daily Rover@Home sessions, these data would imply Rover@Home is a meaningful interaction for the dog and an effective social substitute.

Although the high degree of measurability of this evaluation method makes is powerful, it is difficult to understand how it will generalize to dogs without quantifiable behavior difficulties. Elimination of pathological behaviors does not necessarily imply healthy dogs will similarly benefit.

Finally, there is the ethical component of asking a dog owner to continue a pattern of socially isolating their dog in order to attempt an experimental treatment.

5.2 Clinical Trials

We performed two types of clinical trials. The first were several sessions with Sydney, a 3 /12 year old neutered male Silky Terrier and his owner Bruce. The second type of trials was one-day trials at Gemini Dog training center of Littleton Massachusetts.

5.2.1 Trial 1: Sydney and Bruce

Figure 5-1: Bruce(left) using Rover@Home with Sydney(right). Sydney is "touching" the alley-oop in response to Bruce's vocal command. Note the feeder on top of the red box in the upper left corner of the picture. The webcam is placed on top of the feeder.

Sydney came to Rover@Home with 2 ½ years of clicker-training experience. We performed four sessions in Sydney's home over the course of nine months. All trials used two computers directly connected by a local area 10 Mbit Ethernet connection. For all trials, Bruce was in one room and Sydney was in another. Normal conversation and vocal commands could be heard at an attenuated level from one room to the other. Because of network delay, Bruce's verbal commands would be produced by the PC speaker with about ¼ second delay. The rooms were totally optically isolated – Sydney's only sense of Bruce's presence was audio (and perhaps smell). Sydney had been clicker-trained in this room prior to Rover@Home, so he was not only familiar with the surroundings, but also familiar with being clicker-trained in these surroundings.

Unless otherwise noted, only Bruce and Sydney were present at these sessions.

5.2.1.1 Session 1: October 2000

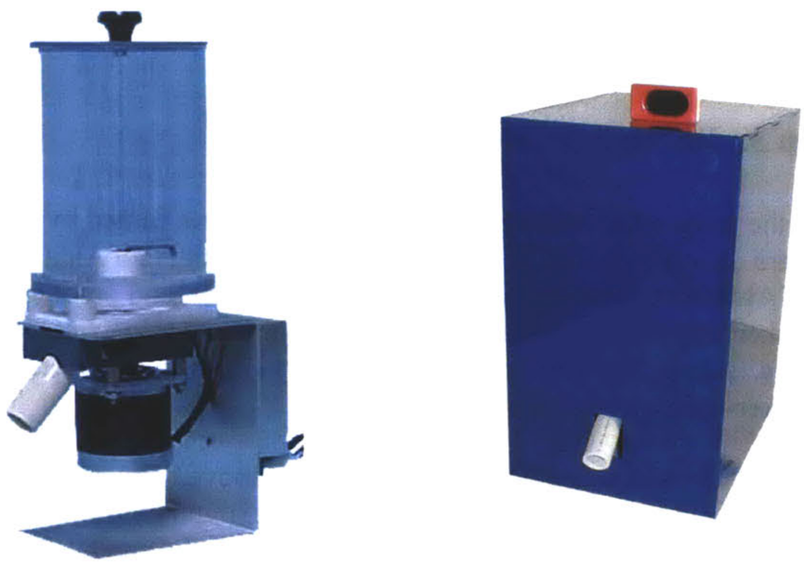


Figure 5-2: Med-Associates feeder (left) and enclosed in tamper-resistant case (right). Use was discontinued because of large size and stutter noise frightened dog. Shown with clicker for size comparison.

This inaugural session used a commercial feeder from Med-Associates encased in a large blue tamper-resistant container (*Figure 5-2*). The largeness of the feeder, coupled with the stuttering and grinding noise it made while emitting food startled Sydney and the sessions were discontinued.

This early version of Rover@Home also included a light to be remotely activated indicating: "session in progress." This was intended to replace some of the sense of presence missing in the remote interaction.

5.2.1.2 Session 2: October 2001

For this session we constructed a custom feeder. It was much smaller in size, used transparent plastic allowing the dog to view the contained food, and made a quieter, more even noise. Sydney was accustomed to the feeder over two evenings. Bruce performed normal co-located clicker-training sessions, but used the computer to deliver clicks and treats. This familiarized Sydney with the unique timbre of a PC Speaker rendered clicker sound, and obtaining food not from Bruce, but from the feeder. We also eliminated the light because of technical difficulties.

After a few iterations of clicking and treating, Bruce did some simple tricks such as sit, down, and touch [the alley-oop]. All trick evaluation was visual, through the NetMeeting video window.

The session lasted about 15 minutes. The session was not videotaped or observed by any third party. Bruce reports Sydney responded to the verbal commands and his behavior was consistent with a co-located clicker training session. We considered this trial a success.

5.2.1.3 Session 3: December 2001

This session was the same setup as the previous session, and was filmed by a professional camera crew for Scientific American Frontiers with Alan Alda. The crew consisted of Alda, a cameraman, sound engineer, producer, and production assistant. Researcher Ben Resner was also present. Sydney was told to sit, down, and touch as before. Again, his behavior was reported by Bruce to be consistent with his co-located clicker-training behavior.

We also included the wireless squeeze toy panda pictured above. Bruce made several attempts to lure Sydney to the panda by activating a digital sound chip embedded in the panda, and rewarding as Sydney increased interaction with the panda. This was the first use of acoustic luring. Bruce reports Sydney responded to the sound of the squeeze toy and began to pick it up on cue. The sessions with the panda were discontinued after Sydney repeatedly removed the electronic innards. Given Sydney's general interest in plush toys, it was unclear if Sydney was actually learning any association between the verbal command "panda," or just exhibiting his natural interest in plush toys.

5.2.1.4 Session 4: January 2001

Again, this was the same setup as before, but without the electronic panda. This session was also filmed for archival purposes. The videotape clearly shows Sydney responding to sit, down, and touch commands. Furthermore, Bruce reports Sydney's orienting towards and attending of the feeder was similar to how he responded to a person. Sydney would remain relatively motionless and attentive while awaiting commands, and only moved in response to a vocal command. This session included researcher Ben Resner.

This attentiveness clearly demonstrates Sydney's ability to recognize and respond to vocal cues when coming from outside the room. Sydney was also able to generalize from a real clicker and hand delivered treat to a digitally sampled clicker and machine delivered treat. Aiding this generalization is the fact Sydney is used to being clicker trained in the exact environment where the Rover@Home testing took place.

It is important to reiterate that Sydney could hear Bruce's commands from the other room about ¼ second before they arrived over the speaker.

5.2.2 Trial 2: Gemini Dog Training Center

Figure 5-3: Gemini Dog training sessions. Paul using the interface (left), and Steiff "plunking"[12] the alley-oop (right). Note the feeder behind Steiff's head.

We also installed Rover@Home at Gemini Dog Training Center, Littleton MA. Three professional dog trainers used the system with four dogs over the course of three hours. Both dogs and trainers were videotaped, and trainers were asked to narrate their experiences. The trainers were given a brief introduction to the Rover@Home interface. All sessions started with a few iterations of clicking and treating with the trainer in the room. Beyond this, the sessions were unstructured, allowing the trainers flexibility in improvising an appropriate course of action. All dogs had been previously click-trained.

The dogs were placed in one room, and the trainers in another. The rooms were acoustically isolated at normal conversational levels. Yelling and other loud noises could be heard from one room to the other.

The trainers used a 900 MHz Dell Inspiron and the dogs had a 650 Mhz Sony Viao Laptop. The two computers were directly connected via a 50-foot crossover Ethernet cable. The Rover@Home setup included the feeder, PC speakers, and webcam. The

[12] By convention, "plunking" or "bonking" generally means touch the base of the alley-oop with a paw, whereas "touch" means touch the tip with the nose.

setup also included an alley-oop wired with a tilt-sensor and optical/audio lure placed on the stem, right below the tip. Similar to the setup with Bruce and Sydney, there was approximately a ¼ second audio transmission delay.

Attending the sessions were the trainers mentioned below, researchers Ben Resner and Bruce Blumberg. Also present was Spencer Lynn, a visiting Media Lab Graduate student from the department of Ecology and Evolutionary Biology at University of Arizona. At different points during the sessions other trainers would come and go, but the mentioned trainer drove each session for the entire time.

5.2.2.1 Session 1: Lilly, Boston Terrier, 7-8 years old, spayed female

Trainer: Janet
Lilly was skittish, and the feeder frightened her. She would eat the food from the feeder, but only tentatively, and nervously attending to the feeder the entire time. She became a bit habituated to the feeder after a few cycles of clicking and treating, but never got close to performing tricks. Lilly had exaggerated displacement behavior – she would get startled, run in circles, lick herself, bark, and then eat the dispensed treat crouched in a flight position. Lilly would have required several sessions to feel comfortable with the feeder. Lilly was also the only dog to actually paw the feeder. This behavior suggests that despite her fear, Lilly understood a link between feeder noise and treat delivery.

We made a few attempts to get her to perform with the trainer in the other room, but gave up after it was clear Lilly would require much more time over many sessions to feel comfortable with this new context.

5.2.2.2 Session 2: Elliot, Miniature Black Poodle, 18 months, neutered male

Trainer: Janet
Elliot is a much braver, more curious and animated dog. It only took one or two clicks for Elliot to overcome any fear of accepting treats from this mysterious dispenser. When we attempted remote training sessions, Elliot became much more interested in the feeder than any vocal commands coming out of the speaker. He was too excited about this strange feeder object to attend to any of the stimulus coming through the speaker. He was not considering the PC speaker as a source of information. After a few tries with Elliot, we gave up.

5.2.2.3 Session 3: Steiff, Jack Russell Terrier, 16 months, neutered male

Trainer: Carolyn
Steiff had a particularly solid clicker-trained background, so we were optimistic. But Steiff seemed largely to ignore his owner's voice coming over the speaker. When I

briefly entered the room to change the tape on the camera, Steiff made play overtures with a ball as if he were all alone and nothing else was going on. This behavior is inconsistent with how Steiff would react to a stranger's entrance during a co-located clicker training session.

Steiff had been trained to interact with the Alley-oop in two different ways. He could "touch" which means touching his nose to the ball on the top, or "plunk," which means touch the base with his paw. One of his first commands to "plunk" produced the behavior, but directed at the feeder, not the Alley Oop. Unfortunately we failed to reward the behavior and by the time we realized our error, the behavior had extinguished. It would have made sense to reward the behavior, even if directed at a different object, in order to maintain enthusiasm with the session. The fact Steiff displayed the plunking behavior shows he has the physiological mechanisms to understand vocal commands from the PC speaker, and can at least partially generalize from his co-located training experience to this new remote context. Had we been more prepared for this behavior and rewarded it we may have been able to strengthen it. A big lesson is for trainers to be very prepared to reward anything that's the least bit similar to any behavior the dog already knows.

5.2.2.4 Session 4: Daria, Burmese Mountain Dog, 9 months, intact female

Trainer: Paul

Daria had the mellowest personality. She seemed to balance Lilly's skittishness and Elliot's kineticism. Daria's owner Paul did not start right off with drills. Instead, he would click, treat, and then say: "find the biscuit" or "find the cookie." His command to the dog was very simple -- obtain the treat from the dispenser. Instead of using the treat as a reward for another behavior, the treat was the reward for finding the treat.

Paul talked about "charging the speaker." Clicker trainers talk about "charging the clicker" which means doing several rapid sequences of clicking and treating. This builds up (charges) the association between clicking and treating, and highlights the clicker as highly relevant behavioral stimulus. Paul was doing the same thing with the computer speaker. Because NetMeeting uses various compression schemes to optimize audio transmission, combined with the fact the PC speakers do not have a full frequency response, what the dog hears may not sound at all like the owner's voice. But if the speaker can be "charged," the dog can learn to pay attention to what it hears, and respond as if it was the owner. From our experience with Steiff's plunking behavior, we know the dogs can parse sound information from the PC speaker.

This strategy seemed to be successful. Daria learned to sit and lay down, and did several cycles before losing interest and wandering off. Daria inserted a superstitious half turn before each sit, which shows she was actually learning a new behavior instead of generalizing from an existing one. Daria was also the only dog to orient to the speaker. This indicates attentiveness to the speaker as a source of information about the opportunity for a click and food reward.

Paul (the trainer) was the most comfortable with the Rover@Home setup. He was the only trainer to operate the interface himself. The other two trainers gave vocal commands while one of the researchers (Ben Resner) held the computer mouse and pressed the buttons. This created obvious difficulties in timing. Having the trainer also operate the interface helped synchronize events.

Finally, Paul was the most comfortable talking to his dog through a microphone while being watched by two people and videotaped. The verbal reassurance was likely a positive factor, and helped create a sense of presence for Daria. The other two trainers were more economical with their verbal praise, reflecting their co-located training philosophy.

5.2.2.5 Session 5: Steiff (repeat)

Trainer: Carolyn

Our success with Daria led us to try another session with Steiff, this time charging the speaker before proceeding to more complex tricks. The second trial we did have success with "down", which is his default behavior -- the first trick he learned, and the trick he reverts to when he's not sure what to do. In contrast to Daria, Steiff was applying previously learned behavior to the new context. Steiff's trainer Carolyn said she felt Steiff would have to relearn a few of his tricks before he "got it" that all his previous clicker training applies to this new context. Carolyn was optimistic Steiff would eventually generalize from co-located to Rover@Home.

5.2.2.6 Session 6: Sydney (Silky Terrier).

Trainer: Bruce

This is the same dog / owner pair from the first trials. We wanted to compare Sydney's performance at the dog-training center with how he performed in Bruce's home. This would give us some measure of scientific control. The trial was rushed as Gemini clients were starting to come in to pick up their dogs from the daycare Gemini runs. Sydney largely ignored the device and was distracted by outside noises in the same way as the other dogs. Bruce reports this distraction was in part because Sydney was aware of Bruce being just outside the room and was angry and frustrated at this partial abandonment. Furthermore, even though Sydney was familiar with being clicker-training in the Gemini training facility, he had no experience being trained in that particular office.

In summary, Sydney's behavior was much more consistent with the other four dogs than with how he reacted in his home setting. This tends to indicate the setup at Gemini was more distracting to the dogs than a quieter, more controlled home environment. This was likely significant for all dogs at the Gemini trials.

5.2.3 Gemini Results Conclusions

- The dogs seemed unexcited about the cat food we used as treats (Many dogs view cat food as a treat – this is hopefully not an example of a poorly understood interface). Food is not motivationally significant to dogs that are overly distracted or not hungry.

- The motor sounds seemed to be as much of a conditioned stimulus as the click. For example, Lilly did not orient to the food until she heard the motor sound. The click seemed unimportant – ambient clutter that contained very little useful information. This is not necessarily a problem. As long as some sort of precise means exists to mark the behavior, the clicker training concept is intact. The motor starting is a sufficiently staccato signal to be an effective event marker.

- Although the dogs were curious about the feeder, none of them attacked it or otherwise attempted to break into it. As previously mentioned, only Lilly actually touched the feeder, and only tentatively.

- Although the dogs had a hard time generalizing from their co-located clicker training to remote clicker training, the trainers had a hard time as well. One of the trainers asked a researcher: "when should I click and treat?" This trainer was told: "click when you would normally click, and treat when you would normally treat." It took a few cycles for the trainer to realize the fundamental similarity between co-located clicker training and Rover@Home. Just as the dogs had difficulty generalizing, so did the trainers.

  This shortcoming can be alleviated by allowing the trainers to first use the Rover@Home interface in a co-located setting. This will give them a feel for the interface, and give them a visceral sense of the relationship between clicking a button to reward, and the sound of the clicker and dispensing of treat. Without this co-located experience, it was perhaps too abstract to click on a button to reward their dog. Bruce (Sydney's owner) had the advantage of seeing Rover@Home working in a co-located setting[13], so he had solid reinforcement for the relationship between abstract mouse clicks and canine rewards.

- The PC speakers were originally placed about two feet apart for stereo separation. We realized this probably means nothing to a dog, and unplugged one of the speakers turning the sound into more of a point source. This seemed to help the dogs orient towards the speaker.

  Additionally, inexpensive PC speakers were used at Gemini, whereas

[13] Bruce supervised the writing of this thesis, so was quite familiar with the development of Rover@Home.

expensive ones were used at Bruce's home. Future trials should investigate the importance of speaker fidelity for Rover@Home remote clicker-training sessions.

- All the trainers talked to the dog as if the dog was aware of the source of the camera. They would say things like: "get back in the picture" or "I can see your tail." At one point a Steiff went off-camera and then suddenly appeared right in front of the camera. Everyone knew this was accidental, but it was funny nevertheless. While camera placement does not matter for the dog, good camera placement seems to not only increase enjoyment for the owner, but also be an important factor in an owner's ability to determine trick compliance as well as dog attention. Camera placement on top of the feeder is probably the best place, because this is a natural place for dogs to orient.

  Optimal camera placement was problematic at Gemini. The layout of the room was such that no one location was both accessible and provided a view of the entire room. This contrasts to Bruce's home where the camera was placed against a well in such a way Sydney would naturally orient towards the camera and could not easily go out of view.

  Additional solutions are to employ a movable camera or multiple cameras. While technically feasible, we feel this is an unnecessary complexity. Additionally, the motorized sound of a moving camera could potentially distract and / or frighten the dog.

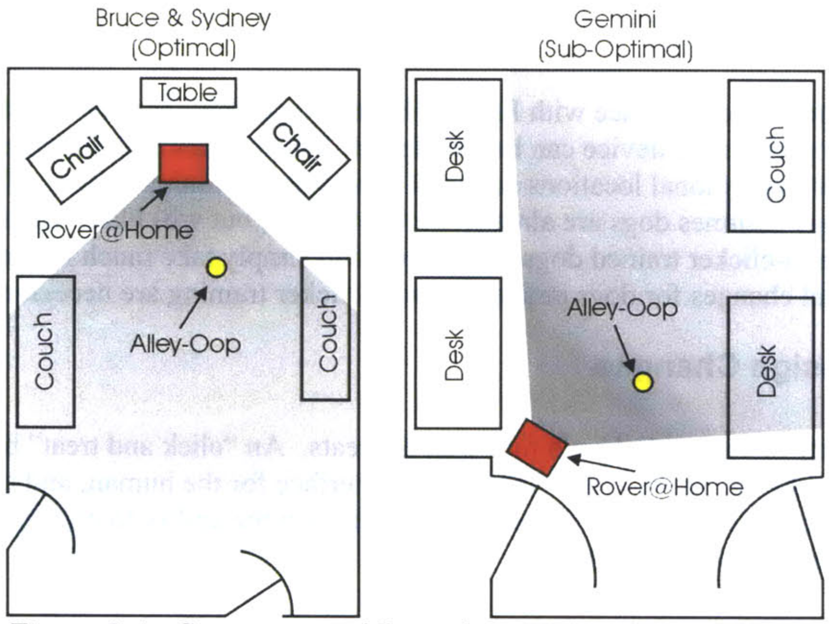


Figure 5-4: Comparison of floor plans at Bruce's home (left) with Gemini (right). At Bruce's home, the camera is placed such that Sydney will not generally be out of view. This contrasts Gemini, where the dogs often went out of camera view to investigate noises coming from the doors.

View frustum of camera is designated in gray.

- One of the trainers suggested including a punisher such as an automated sprayer. The owner could release treats and then punish the dog for eating the treat without waiting for a verbal "eat" command. This would force the dog to attend to the speaker as an information source about when food can be consumed. This could also extinguish any interest in Rover@Home.

6 Discussion

While actual testing was brief and informal, it tends to support our hypothesis that Rover@Home successfully reproduces enough of the clicker-training interaction to be a similarly rewarding experience for owners. Owners report their dogs either respond favorably to Rover@Home, or are confident they would if given sufficient exposure. None of the dog owners expressed reluctance towards using the device, or in any way felt Rover@Home posed a hazard to their dog.

An unexpected result was the need for increased emphasis on owner training, and giving the owner time to become familiar with the interface. Just as the dog needs time to generalize from co-located to Rover@Home, so does the trainer.

6.1 Future Trials

From our present experience with Rover@Home, a methodology for introducing new dog/owner pairs to the device can be developed. This applies to both in-home trials and trials at institutional locations such as Gemini Dog Training Center. This methodology assumes dogs are already clicker trained, but will likely work just as well with non-clicker trained dogs. Each step will simply take much longer. But no fundamental changes for dogs unfamiliar with clicker training are necessary.

6.1.1 Design Changes

In a majority of cases, clicks are followed by treats. An "click and treat" button that does both with one press would simplify the interface for the human, and perhaps optimize timing for the dog. We still give the human the ability to click without treating and treat without clicking.

Furthermore, in light of the fact dogs seem to attend to the starting of the motor as the conditioned stimulus, collapsing the click and treat into one signal could provide clarity for the dog. Future trials should investigate this change as well as eliminating the clicker entirely and only using the start of the motor as the conditioned stimulus. In this case, it would be nice to have some means of starting and stopping the motor without actually releasing a treat so trainers can offer the conditioned stimulus without the associated treat.

6.1.2 Setup

Spend some time investigating the dog's favorite feeder-compatible treats. Place the feeder in corner of room with camera on top of feeder. The top of the feeder is the ideal vantage point for the human, and placement in a corner makes sure dog cannot go off-camera. Use only one speaker placed close to the feeder. With this arrangement, the speaker, feeder, and camera are all very close. Dogs attending to either the speaker or the feeder appear to be gazing at the owner.

6.1.3 "Charge" the Feeder

Do some co-located clicks and treats with the feeder until dog understands clicks/motor starting is an indicator of treat. Do this until dog is no longer startled at feeder. Could potentially take several sessions and much trainer reassurance. The feeder is charged when the dog orients towards the feeder in response to a click.

6.1.4 Do some co-located clicker training

In a co-located setting, run through some of the dog's strongest tricks with clicks and treats coming from the computer instead of the trainer. This will also accustom the trainer to the computer interface, and reinforce the trainer's ability to apply their clicker training acuity to the remote setting. It's not just the dog that needs to generalize.

As in co-located clicker-training, it is important for the trainer to keep the dog interested in the training session through a reward schedule that motivates the dog to work.

6.1.5 "Charge" the Speaker

This is the first remote dog-human interaction. Get the dog used to hearing a voice coming from the speaker. This trains the dog that sounds coming from the speaker indicates the opportunity to perform a behavior that can earn a reward. Charging the speaker consists of the following three steps:

(a) say: "find the treat" / "find the cookie" (or similar enticement)
(b) click
(c) treat

This is a different order from how Paul "charged the speaker" (he did click-treat-speak). By placing the verbal command first, it initiates a series of events that eventually leads to a treat. Even though Paul did have success, we feel this arrangement will be even better. Stimuli occurring after the unconditioned stimulus are less motivationally significant.

The transition from co-located to remote interaction can optionally be broken down into smaller steps where the owner is not completely acoustically or visually isolated from the dog. Recall that the sessions with Bruce and Sydney did not have complete acoustical isolation. This could have been a factor in Sydney's relative ease to generalize from clicker training to Rover@Home.

6.1.6 Train behaviors

Start with the dog's strongest tricks. Trainers should do their best to talk to their dog continuously in order to maintain a presence. Provided the speaker has been charged, dogs should already be attentive to the PC speaker.

Trainers should be prepared to reward incomplete or alternative versions of existing behavior in the interest of sustaining motivation. But this is not just temporarily lowering expectations, it entails being open to an alternative set of expectations. Had we rewarded Steiff's blonking behavior directed at the feeder instead of the ally-oop, we may have been able to shape a behavior and sustain his interest in Rover@Home. We were too focused on the actual behavior to spend time creating enthusiasm about the Rover@Home device.

6.1.7 Project Owner Presence

We should experiment with other means of projecting owner presence beyond constant vocalization. In some environments, constant chattering into a computer microphone might not be appropriate. Reintroducing the training light used in the earliest working version could be sufficient. We would expect optimal results when the light is also illuminated during co-located clicker-training sessions. This reinforces its meaning as "training in progress now – opportunity for human interaction and treats."

Another possible means to project owner presence is scent. As discussed earlier, the best "scent renderers" are about equivalent to a tape recorder that is either off or repeating "I'm thinking of you … I'm thinking of you." This is all the information necessary to project a sense of presence. If a particular odor is always associated with co-located clicker-training sessions, that odor could serve as a powerful cue towards generalization for remote clicker-training. Because odor is unused in all other parts of the Rover@Home interaction, it presents an uncluttered sensory space for remote digital control.

6.2 Applying HCI to Animals – Three "Gotchas"

During the Rover@Home design process, mistakes tended to be combinations of three issues, which we discuss below. These gotchas are certainly a problem with traditional human-computer design, especially with children, but are particularly difficult to avoid when designing for animals. All three of these gotchas are variants of careless anthropomorphism.

Anthropomorphism is misleading not just to interface designers, but animal trainers as well. For trainers, anthropomorphism leads to making excuses for the animal, and reinforcing for projected beliefs of the animal, not for actual behavior (Ramirez, 1999).

6.2.1 Gotcha 1: Giving Human Technology to Animals

Interfaces designers for animals need to be careful not to take devices built for humans and simply put them in front of animals. One needs to respect the fact that tools appropriate in a human task domain are not necessarily useful in a non-human task domain. When searching for a remote interaction between dogs and humans, video was an intuitive and almost obvious first step. Nonverbal communication immediately makes most humans think of video. This is almost certainly influenced by the wide availability of video rendering devices for humans. The existence of high-resolution information rich display devices creates an incredible temptation to include video in an interface designed for an animal with visual acuity in the same order of magnitude as a human. It was only after we examined canine physiology and psychology that we realized present-day video technology would probably not be a cornerstone of a remote interaction.

To some degree use of human technology is unavoidable. Development of similarly sophisticated display devices optimized for animal vision is well outside the scope of this thesis. If it were felt that some type of electronically reconfigurable display was crucial, we would probably just have to make do with a traditional CRT or LCD display and hope to optimize the interaction to make use of the strengths of the display with respect to the creature's ocular physiology.

6.2.2 Gotcha 2: Devices That Lead to Humanlike Interactions

This usually manifests itself as the temptation to build devices that would lead to funny or cute interactions from a human perspective. A common comic technique is for an entity to behave in a manner beyond its accepted mental and physical capabilities. Television commercials are filled with toddlers talking about their brokerage accounts or environmentally conscious performance sedans. People find this entertaining because kids this age do not normally discuss adult matters in such a matter-of-fact tone of voice. Other commercials show animals driving a car or using a credit card to buy mailorder dog toys.

"On the Internet, nobody knows you're a dog."

Figure 6-1: Everyone knows dogs do not use the Internet, yet this now famous cartoon has a powerful influence on design sensibilities. When we describe our work, this is what people think.

When we describe our research, we often hear: "Wouldn't it be cool if your dog could send email to its owner or surf the Internet." This would be entertaining to humans, and would further a desire to elevate animals to human status, but it is inconsistent with what animals want to do. The device needs to be good for the dog.

It needs to deliver something the dog actually desires, not a projection of our own desires. While a dog could potentially alert its owner of its desire to play by activating a device that alerts the owner through an email, it is not really fair to claim the dog is "using the Internet" in any way similar to how most humans "use the Internet."

Owners can even go so far as to train a dog to use a device in such a way that is amusing to a human. But that does not mean "the dog likes it." The dog may seek out the interaction in the hopes of being rewarded, either with a treat, or with increased social interaction. But all that has been shown is the dog is physically capable of using such a device. There is nothing intrinsic in the device or interaction that brings pleasure to the dog. It is simply a prop that mediates human-dog interaction. Creation of an artifact a dog enjoys is certainly not a bad thing, but one needs to be careful to understand the true motivation behind any interaction, especially when attempting to learn from the interaction in order to make to more general or increase its appeal.

When searching for a suitable interaction between animals and computers, it took us a while to realize clicker training was the perfect established ritual. Clicker training has no real analogue in human-human interaction[14] and it does not involve fancy gadgets. It was not immediately obvious that a 50-cent clicker held the key to remote interaction between humans and dogs. It was only when we looked carefully at dog-human interactions as distinct from human-human interactions that we realized the power of clicker training.

6.2.3 Gotcha 3: Devices Uninteresting to Humans

It is tempting to think that because animals are incapable of performing many human tasks, they are similarly incapable of appreciating all human sophistication. This can lead to "designing-down" to the perceived inferiority the animals. If a device is uninteresting to a human, one needs carefully to consider why it might be interesting to an animal. For example, if a device produces an overly repetitive output, one should question why an animal would not become bored of this same output when a human most certainly would.

When we developed a music switcher for "InterPet Explorer," an interactive electronic environment for African Grey Parrots, Alex, a 25-yearl old bird, was initially using the device, but then his interaction dropped off after a few days. We realized his only options were the same four songs, over and over and over. No wonder he became uninterested – without fresh content, he became bored. But we did not realize this shortcoming until we evaluated the device from our human perspective. Somehow, we thought that because the device is intended for a bird, it

[14] Operant conditioning, upon which clicker training is based, certainly exists among humans. But we know of no human-human rituals that explicitly depend upon operant conditioning. In fact, explicit behavioral control of children with operant conditioning is generally frowned upon.

would be interesting in some way totally unfathomable to humans. This shortcoming was quickly remedied with the addition of live Internet radio streams.

However, one needs to be open to the possibility of the existence of a device that is captivating to animals, but quite dull to humans. In the same way the television series "Tele-Tubbies" is fascinating to one year olds, yet a complete mystery to most sober adults, one need to be on the lookout for an interaction that might captivate our pets even though it bores us. Being uninteresting to humans does not mean a device will similarly be uninteresting to animals – but it could be, and for the same reasons. The domain of interactions uninteresting to humans but compelling to dogs is the hardest place to look, and is perhaps the most fruitful. As mentioned in the last section, clicker-training was not immediately obvious as the basis of a remote human-animal interaction.

6.3 Analysis of Rover@Home as Co-Located Substitute

Steve Benford et al analyze remote interactive environments according to three variables: Transportation, Artificiality, and Spatiality (Benford, Greenhalgh et al., 1998). We find this a useful framework in understanding Rover@Home and possible optimizations. In our design of an asymmetric interface for human-animal communication, we come up with similarly asymmetric values for transportation, artificiality, and spatiality.

Transportation is the degree to which the participants leave their local space and enter a new remote space. With Rover@Home the human leaves their office computer and enters the world of their dog. The human's focus is on the video feed from the location where their dog resides, not on their local surroundings. The owner sees the familiar furnishings of the dog's present surroundings, and is transported there. For the dog, there is no transportation. The dog remains in his home, and is given no sense of the owner's setting. The dog has no indication of the owner's location.

Artificiality is the degree to which the shared environment is synthetic. Immersive virtual reality is highly synthetic in that the world rendered to the user is totally the creation of the computer. For the human, Rover@Home is very unsynthetic. The environment the user observes is not computer generated, but most likely the owner's home. For the dog, however, the setting is more artificial. Even though the objects in the dog's world are real, the treat dispenser and computer-generated clicker sound are synthetic elements compared to a live owner with a mechanical clicker. To a dog, these artifacts are substitutions for human presence and therefore artificial. Just as humans view a VRML (Virtual Reality Markup Language) 3D world as a synthetic representation of the real world, the dog will probably respond to the feeder and clicker sounds as variants of more familiar objects, and its behavior towards these objects reflects non-synthetic experience.

Spatiality is the fidelity with which the medium supports the ability to express Cartesian relationships such as containment and orientation. The human observes the dog's surroundings and can therefore evaluate the dog's location in its environment. To a lesser degree the dog has a sense of the owner's spatiality as expressed through interaction with the instrumented props such as the alley-oop and squeeze toy. The owner can lure the dog to a specific prop by activating its sounds or lights. By changes in the dog's physical location or gaze through activation of a prop, the owner can project presence in different parts of the room. Spatiality for the dog could easily be enhanced with a computer controlled laser pointer accompanying commands such as "over there" or "touch laser."

In terms of audio, the dog has no sense of the owner's spatiality. The tested version of Rover@Home uses a single speaker and no video, yielding no spatial information. The owner is completely unable to augment commands such as "over there" with spatial cues such as pointing or audio orienting. This begs the question of how increased acoustic and visual spatiality provided by stereo speakers enhance the dog's participation in Rover@Home.

6.4 Is Rover@Home a "Tangible Interface?"

A common misperception of tangible interfaces is to think anything attached to a computer that is not a keyboard or mouse is a tangible interface. But there is an emerging intellectual framework for evaluating tangible interfaces (Ullmer and Ishii, 2000), and it is instructive to see how Rover@Home fits in this paradigm.

6.4.1 TUI Framework

Below we analyze Rover@Home according the four principles for Tangible User Interfaces as described in "Emerging Frameworks for Tangible User Interfaces." (Ullmer and Ishii, 2000).

6.4.1.1 Physical representations are computationally coupled to underlying digital information.

When the owner clicks the "treat" button on the interface, an electronic ping is sent to the dog's computer. This causes the computer to release a treat for the dog. If the dispensed morsel of food is considered part of the interface, then Rover@Home presents a clear example of digital information (electronic ping) being rendered through a physical device (food morsel). We propose this is the first example of an "Edible User Interface (EUI)," which we introduce as a legitimate subset of the tangible variety. The owner's action is clicking a button, but the dog's perception of this action is receiving a treat.

6.4.1.2 Physical representations embody mechanisms for interactive control.

Rover@Home uses very different mechanisms for input and output. The treat dispenser, click sound, and speaker are purely output devices. The dog earns treats through compliance with verbal commands, not by any direct manipulation of the treat dispenser or speakers. In fact, a typical Rover@Home user will likely discourage tampering with the dispenser by the dog. Therefore the physical representations of input and output are quite distinct.

The wired alley-oops are perhaps closer to providing interactive control because they have input and output sensors on the same device. But the inputs (tilt) and outputs (flashing light / sound) are strictly either input or output modes. The lights flashing does not change the nature of how the alley-oop inputs information; similarly, the act of being tilted does not change how the device flashes and beeps. With the current alley-oop, there is no means of interacting that treats it simultaneously as an input and output device.

6.4.1.3 Physical representations are perceptually coupled to actively mediated digital representations.

Let us consider the owner's vocal commands over the PC speaker to be a digital representation of the owner, or at least the owner's intention. Generally, when humans think of "digitally reconfigurable representation," they think of video displays, but because there is no video display for the dog, electronic reconfiguration takes the form of a PC speaker. As pointed out earlier, we consider the treat to be a physical representation of owner satisfaction. The click can actually fall in either category. It is rendered by the speaker, but it is strongly linked to a physical device. (in fact, a test version of Rover@Home used a solenoid to activate an actual clicker – from the dog's perspective there is no difference between the two). The owner gives the dog a command, and if the dog's behavior is satisfactory, the dog gets a click and a physical treat.

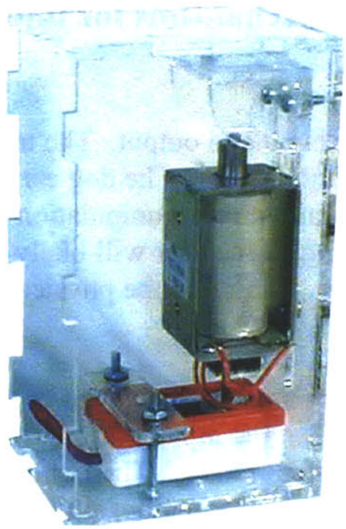


Figure 6-2: The mechanical clicker originally used in Rover@Home. The solenoid presses down on the orange clicker. Because of the different resonance of the box, this clicker sounded different than when held by a user's hand. For this, as well as reasons of simplicity, we decided to instead render the clicker sound through the PC Speaker.

The association between the physical representation (treat) and digital representation (owner's voice) is greater still if the treat is also accompanied by verbal praise. This strongly couples a physical morsel of food with a digitally reconfigurable acoustic rendering of the owner's voice.

For Rover@Home to be effective, both the physical and digital channels are necessary. The owner's disembodied voice without any reinforcing stimulus does not provide the dog with a channel to respond, and a treat dispenser without reconfigurable acoustic rendering does not allow much variation in the interactivity.

6.4.1.4 Physical state of tangibles embodies key aspects of the digital state of a system.

If we consider the dispensed food part of the interface, then each morsel consumed by the dog represents a physical change in the dog's hunger. This in turn can affect the motivational significance of additional treats. As the Rover@Home interaction progresses, the dog becomes increasingly satiated, a physiological effect that alters the state of the interaction. By measuring the dog's hunger or motivation to earn an additional treat, a user can infer the duration of the Rover@Home session and the success of the dog's ability to obey verbal commands.

This line of reasoning is open to debate. One could use this logic to say a computer monitor is tangible because it causes eye fatigue, which can be measured and therefore used to infer the amount of time the user has been staring at the screen. Or one could measure tendon strain to gauge keyboard use. These are both instances of digital information (duration of exposure to device) encoded in physical state (fatigue).

We feel the concreteness of the food combined with the tangibility of caloric intake separate Rover@Home from the above reasoning by a matter of degree. Agreed, there is no fundamental difference between the two. But the caloric intake represents an intentional and deliberate change in the system, not an unintended consequence of use. Keyboards and monitors are not designed to cause strain; this is simply a byproduct of their use.

6.4.2 TUI Analysis

While there is a strong case for Rover@Home being a TUI, it falls short of full inclusion in the TUI framework. If the basis of the interaction were the dog manipulating some type of device that could release food, this would bring Rover@Home closer into the TUI realm. In this scenario, the means of control (dog correctly manipulating device) are closely coupled to the physical representation (release of food).

Not to worry. Our goal for Rover@Home was not to design a TUI for a dog. We analyze Rover@Home as a TUI simply to understand the relationship between the two, and the hope it will point us to design optimizations.

6.5 Rover@Home as a Ubiquitous Interface

From the dog's perspective, Rover@Home is much closer to being ubiquitous (Weiser, 1991) than tangible. In his now landmark 1991 article, Marc Weiser predicts the trend in computing towards machines seamlessly integrated into our environment. The dog is certainly unaware of the existence of a computer mediating the interaction. While the interaction may be strange to the dog, the dog is unaware where the computer ends and the real world starts. For the dog, it is all real world. The computer is nicely hidden well outside the dog's awareness. The feeder and other Rover@Home props are special-purpose devices designed to imitate, as closely as possible, the real-world analogues the dog has become accustomed to. Furthermore, Rover@Home allows the dog to perform a specific task – participate in a clicker-training session.

Adapting computers for animals will probably follow this strategy. General-purpose configurable devices will probably be too difficult for a non-human to learn (they're difficult enough for many humans to learn). Single function devices designed for a specific implementation are much more suited for the use by animals.

7 Applications

7.1 Operant Conditioned Interfaces

J.R. Anderson suggests that learning unfolds in three stages (Preece, 1994)[15]:

[15] From J.R. Anderson: The Architecture of Cognition, Harvard University Press.

1. Cognitive Stage: Acquiring declarative knowledge relevant to the skill being learned. This includes memorizing subtasks important to the overall goal

2. Associative Stage: Integrating various bits of knowledge learned in the cognitive stage into larger chunks. This is also where many errors are eliminated

3. Autonomous Stage: The declarative knowledge of the previous two stages become secondary, and the skill is reflexively performed without active cognition

Rover@Home goes through these stages in reverse. We start with the autonomous conditioning of an association between food and a click. Only after the dog has a firm association between clicking and treating, does a user move to the more cognitive elements of the interaction – the dog figuring out what trick the owner desires. Clicker trainers talk about the animal "getting it," when they figure out that clicker training is a guessing game of eliciting the desired behavior (Pryor, 1984). Dogs come to understand that clicker-training follows a higher-level form where the owner wants the dog to do something, and the dog has to figure it out – and be rewarded in the process.

Wendy Macaky et al created a human controlled interactive character named "McPie" used to shape human behaviors (Mackay, Svendsen et al., 2001). For this installation, an unsuspecting user and the character would interact, with a hidden researcher controlling the character. The character attempted to get the user to touch his hands above his head. When the user made hand motions above his or her head, the character would smile and gaze at the user. When the user made motions away from his head, the character would look away and frown. In this way, the user was encouraged to perform the desired behavior.

Similar to how Rover@Home builds its foundation on principles of operant conditioning the McPie project is an example of applying these same principles to humans. We suggest this is a fruitful research direction. Instead of exclusive reliance on conscious cognition as a means for training users in new interfaces, operant drills could be an efficient replacement to reading user manuals most users do not read anyway. For instance, key combinations or simple operations could be operantly trained before any explanation of the interaction is explained.

For example, the Palm Pilot taught its custom "graffiti" pen input language not through charts and diagrams, but by creating a game that gave users the opportunity to practice keystroke entry in a time-pressure situation.

7.2 Applicability to Very Young Children

Dogs are similar to very young children in that they have a difficult time expressing themselves verbally to adults -- telephone mediated interactions with babies are generally one-sided and unsatisfying. Just as Rover@Home as enabled interaction with dogs through non-verbal channels, it could point the way towards remote interaction between very young children and traveling parents.

From the perspective of remote interaction, the main difference between dogs and human babies is that dogs are typically left unsupervised during the day, while pre-verbal children have constant supervision. The expectation of adult intervention while the baby is interacting with the remote parent could dramatically alter the nature of the interaction, and create exchanges not possible in an unsupervised setting.

The methodology developed here is applicable to small children -- search for a ritualized interaction between parents and children that is amenable to transmission over a digital network. As with dogs, this mostly eliminates touch, and will likely emphasize sight, sound, and interaction with instrumented props.

Similar to Rover@Home, a remote interaction with a very young child could very likely center around some type of training or instruction. Parents spend a great deal of time teaching their children labels such colors, shapes, animal sounds. There is no reason for this interaction to be restricted to co-located settings. Instrumented children's toys containing shapes and colors could be activated by a parent in connection with verbal reinforcement about object properties.

7.3 Consumer Product

Perhaps the most obvious application of Rover@Home would be a commercial product allowing consumers to interact with their dog while geographically separated. As stated in the introduction, the pet market is strong and growing, and pet owners have a strong desire to interact with their dogs while separated.

From a technical perspective, Rover@Home faces similar hurdles to those of the home automation industry[16]. Long a science-fiction dream, the widespread deployment of highspeed Internet connections is making home automation once again attractive to entrepreneurs and inventors alike. Similar to home automation, Rover@Home in a commercial setting faces issues of security and reliability. Installing software and hardware that allows visual access into one's home requires a strong measure of trust. As far as home automation is concerned, once a connection from an arbitrary computer to one's home has been made, using Rover@Home to interact with one's dog is technically similar to controlling and monitoring any homebound appliance. Therefore, Rover@Home makes an excellent add-on to differentiate otherwise similar home-automation services.

Rover@Home requires a broadband connection at both work and home. Of the 130 million U.S. Internet users (Gannett News Services, 2001), 33% have access to a

[16] http://www.xanboo.com

broadband connection, either at work or at home (Arbitron Inc & Coleman Associates, 2001). Of this broadband connected population, 15% have broadband at both work and home (Arbitron Inc & Coleman Associates, 2001). This translates into 6.5 million homes with the infrastructure necessary to run Rover@Home. Furthermore, an additional 2.8 million people have broadband at home, but not at work (Kinetic Strategies, 2001), making them potential users of Rover@Home in remote environments other than the workplace. It is unclear how broadband connectivity and 60% of dog owning households overlap. A linear overlay would yeild 3.9 million dog owners capable of running Rover@Home.

7.4 Pets and Health

There is a strong link between pet ownership and health (Beck and Katcher, 1996). Friedmann et al have shown increased survival rates for pet owners one year after a coronary care unit (Friedmann and Thomas, 1995). Similar studies have increased survival rates for other ailments. It follows to reason this effect would only increase if patients could interact with their dogs while still in the hospital, not just after they get back home. Although hospitalized patients can receive visits from family and friends, they are totally cut off from their pets. This is especially troublesome for people living alone, whose pets experience tremendous isolation during owner hospitalizations. Rover@Home could at least partially alleviate this additional stress by allowing hospitalized pet owners to continue interacting with their pets. This could be especially beneficial for people undergoing longer-term hospitalizations or placement in a retirement home.

It is unclear how pets have this beneficial effect on health. One theory is that pets increase the amount of social interaction for their owners (Beck and Katcher, 1996). Presence of a dog was the largest factor in increasing the number of random social interactions to a test subject. Changing clothing from scruffy to neat, shaving, or gender had smaller effects on the number of random social interactions than presence of a dog (McNicholas and Collis, 2000).

Following this logic, Rover@Home would not have the same health benefits because it does not engender public social interaction. Therefore Rover@Home is an ideal research tool for teasing out the mechanism by which pets have such an incredibly positive effect on humans.

7.5 Pets and Disabled / Elderly

Studies have shown connections between the social acceptance of disabled people and pet ownership. Rover@Home can be viewed as an assistive device allowing physically disabled people to have interactions with their dog. Someone unable to operate a mechanical clicker or throw a dog treat could instead use any number of existing access devices such as eye position sensors, tongue controls, or voice interfaces to connect with Rover@Home. This could enable pet ownership for a population that previously lacked the means to interact. This could also strengthen the bond between disabled persons and their service dogs. Finally, Rover@Home could be a powerful tool for a disabled person to further optimize the utility of a service dog's behavior.

8 Future Work:

8.1 Other Animals

There is nothing unique to dogs about clicker training. Clicker training is based on applied operant conditioning, which is universal to all animals, including humans (Gray, 1994). Any animal motivated by treats dispensable by the Rover@Home treat dispenser and capable of hearing verbal commands over a PC speaker can theoretically interact with a human over the Internet using Rover@Home.

Clicker training was chosen for dogs in part because of the growing number of clicker trained dogs. This represents a large "installed base" of users ready to use Rover@Home. The success of clicker training with dogs demonstrates the real-world practicality of using clicker training as the basis of a remote interaction.

In practice, some animals may not respond as dogs to Rover@Home. Some animals may never overcome fear of the feeder. Others are less food motivated or have a very short attention span.

When generalizing from dogs to other animals, perhaps the most obvious choice would be dog's closest genetic cousin, the wolf. Dogs are descended from wolves, and share the same number of chromosomes. Ray Coppinger's theory of neotoany states that dogs are simply wolves frozen in a juvenile state of playfulness, exploration, and subservience to a master. Similarly, domestic dog morphology of shortened noses and floppy ears corresponds to the juvenile wolf. This would imply wolf puppies have the most in common with dogs, and a Rover@Home setup with wolf puppies would have the highest chance of success.

But Coppinger's neotany can be interpreted another way. Perhaps what is relevant is not the genetic ancestry, but the process of domestication. It could turn out to be that domestication is more significant than lineage. *Canis familiaris* (domesticated dog) could have more in common with *felis domesticus* (domesticated cat) than with *canis lupus* (gray wolf).

Domestication is the process of selective breeding to make an animal compatible with human lifestyle[17]. But animals sharing domesticity share more than just a close relationship with humans. Dmitryk Belyaev selectively bred successive generations

---

[17] A reasonable argument can be made to include humans as an instance of domesticated animals. Some might consider this to be self-referential and circular, but Steven Jay Gould has referred to humans as "neotinized apes", and Scott and Fuller have speculated that not only are humans domesticated, but the evolutionary future of humans will closely follow the evolutionary history of dogs (Scott, J. P. and Fuller, J. L., 1965).

of wild silver foxes, selecting only for tameness. In the process of domesticating these foxes for 20 years, they began to demonstrate many features displayed in domesticated dogs. For instance, their coats became piebald (Spotted or patched coloring) and they had floppy ears (Copinger and Schneider, 1995). These traits were not selected for – they did not even exist in their wild analogues. They simply appeared as a byproduct of domestication.

Domesticated animals, particularly pets, are social animals that crave human interaction. There is no reason to think a bear would enjoy a remote interaction with a human when wild bears (and humans) generally retreat in fear. Furthermore, because Rover@Home was developed by using human-centric HCI methodology, the more integrated an animal is into human lifestyle, the more relevant this body of research is likely to be.

When generalizing this work, it is important to distinguish between generalizing Rover@Home specifically, and the underlying methodology used to search for established rituals and appropriate an interaction suitable for remote computer mediation. Outlined below are two possible interactions modeled on the methodology, not a speculation on how these animals would respond to Rover@Home.

8.1.1 Cats

Cats are actually more popular a pet in the United States than dogs, with 74 million cats versus 61 million dogs (American Pet Association, 1998). While Rover@Home is theoretically useful to cats, human-cat interaction does not normally center around training in the same way it does with dogs. There are two popular conceptions about cats that make clicker training much less popular – cats are untrainable, and cats do not need to be trained in the first place (Wilkes, 2001). It therefore makes sense to back up a step and see if we can apply our design methodology to other ritualized human-cat interactions.

Perhaps the most ritualized low-contact cat-human interaction is the hunting instinct. Just as most every dog owner knows how to play fetch, most every cat owners can successfully interact with a cat by pulling a piece string along the floor. A computer interface that would allow the owner to manipulate a piece of string or other moving target for the cat would perhaps elicit the same predator behavior as when the owner is locally responsible for the motion.

To what degree will this interaction feel like interacting with the cat, or simply remotely controlling a wired toy that interests the cat? We did not specifically ask Rover@Home users questions about their interaction, but they tended to communicate with their dogs as if there were interacting directly with them, albeit at a lower fidelity. With connected cat toys, will owners feel as if they are truly playing with their cats, or simply manipulating an object in which their cat shows interest?

8.1.2 African Grey parrots

Parrots are perhaps an exception to the universality of clicker training. Because of their incredible vocal ability, they learn to imitate the clicker sound and can therefore self-reinforce (Farlow, 2001). Substituting some type of non-imitatable conditioned stimulus such as a light might make Rover@Home appropriate for the birds.

We have had the opportunity to work with Alex, a 25 year old African Grey parrot who is the research bird of Dr. Irene Pepperberg. Alex has been taught a functional 100 word vocabulary through a unique training method called "model/rival training" (Pepperberg, 1999). Similar to clicker training, M/R training does not require touch and is amenable to online interaction. With M/R training, two trainers take turns asking questions about object properties and modeling the correct answer. But the trainers not only model the correct answer, they are also rivals for the bird's attention. Correct answers not only earn the object in question, they earn parrot-desirable social rewards such as praise and eye contact.

Model/Rival training is highly ritualized, making it an ideal basis upon which to model a remote interaction. One trainer would be in the room with Alex. The other trainer would be remote, interacting through two-way videoconferencing equipment.

8.2 Breeds and Breed Groups

As stated in section 3, we chose to ignore breed differences. It was argued that Rover@Home is primarily a cognitive activity dependant on the dog's ability to learn and generalize from co-located clicker training sessions.

But nothing should be assumed. Identifying relevant canine subgroupings and evaluating Rover@Home with respect to these groups could yield interesting results.

Furthermore, if we are using this methodology to design a new canine-human interaction, then specialization along breed abilities could play an important role in optimizing an interaction.

8.3 Mutual Awareness

It is important to point out our goal was not mutual awareness, but an actual interaction that occupies foreground attention for both human and dog. A growing body of literature investigates ways in which people can retain a sense of each other without a full interaction (Buxton, 1995; Strong and Gaver, 1996; Pederson and Sokoler, 1997; Tollmar, Junestrand et al., 2000). A dog's gentle breathing, warm heft, or even aroma are all continuous indications of a canine presence. Sensing, transmitting, and reproducing these remotely would recreate that same ambient feeling of dog, even while owner and dog are separated. Writers discuss not only the activities they do with their dog, but also the ongoing companionship and sense of not being alone when in the company of their dog (Budiansky, 2000; Grenier, 2000).

We find this a worthwhile goal, and one that is sure to challenge the emerging methodologies surrounding continuous remote awareness. Furthermore, much of the work developed here will be applicable to that effort, especially the focus on asymmetric interfaces. The sensory stimuli that allow us to maintain an awareness of our dog are likely radically different than how a dog maintains a sense of us.

Furthermore, development of technology to support background awareness can greatly augment a foreground interaction (Buxton, 1995). The two modes are by no means exclusive. Subconscious background information could not only greatly aid the foreground interaction, but also act as powerful cues for the appropriateness of initiating a Rover@Home session. Wouldn't it be great if your dog could take the same co-located signals he or she uses to appropriately time requests for a walk, dinner, or play and bring that to an online setting? What if you could be interrupted by your dog at suitable times for a Rover@Home session? Wouldn't it be great if your dog could remind you to relax and play during the day?

8.4 Playfulness

Rooney et al show that dog-dog play is motivationally distinct from dog-human play. "In an observational study of 402 dog-owner partnerships … the performance of dog-dog play does not seem to suppress the dogs' motivation to play with their owners as would be predicted if they were motivationally interchangeable" (Roony, Bradshaw et al., 2001). One wonders how computer-mediated interactions fit into this framework. Will interaction with Rover@Home affect a dog's desire to play with either a human, another dog, or more Rover@Home? From which "play-bucket" if any, does Rover@Home come?

8.5 Video for Dog?

In the methodology section, we postulated the dog would not benefit from video of its owner. Given the constraints of display technology and a dog's visual perception, one would not expect a dog to be able to understand images on a computer screen to represent an owner's emotions or intentions. But without experimental backup, this remains speculation. How does adding a video channel to the interaction enhance the training experience? Do dogs learn faster or retain interest longer with the inclusion of video, or does it act as nothing more than a bandwidth consuming distraction? Are these effects dependant on the size and type (CRT, LCD) of monitor used?

8.6 Representation of Owner?

How does the dog perceive the owner? When a remote owner says: "Come here," where do dogs go? Do they approach the computer speaker – the source of the sound, or do they approach the source of the food? We purposely placed the speaker and feeder proximate to avoid this issue. How would separating the two affect the interaction?

8.7 Multiple Dogs?

So far we have only experimented with single dogs. What about owners with multiple dogs? How does Rover@Home scale to several dogs? Is it possible to teach turn-taking where each dog serially works on a different trick? Will the dogs possibly learn from each other?

8.8 Multiple Owners?

Similarly, the Rover@Home interface allows multiple owners to control the clicks and treats, but only the first person to log on gets video. Assuming this technical shortcoming can be overcome, what would a 3-way interaction where all participants were remote be like? Is there a way to manage the inevitable confusion that would result? Why would someone want to do this in the first place?

9 Conclusions

Rover@Home is the closest device we have to a telephone for a dog. Rover@Home allows a roaming owner to interact with his or her dog from virtually any Internet enabled Windows computer. In creating Rover@Home we leverage a familiar and established co-located interaction to settings where dog and owner are geographically separated.

Rover@Home shows promise because it adheres to the sound design principles originally developed for humans. In extending interface design to non-human animals, we show that principles such as cognitive modeling, task domains, direct manipulation, and affordances have direct analogues in the non-human realm.

We introduce the concept of asymmetric interfaces that address the unique sensory, cognitive, and motor skills of each creature. In designing an interaction between two different species, we emphasize the importance of catering to both creatures equally and not making one creature adapt to the other, or adopting a "least-common-denominator" approach.

Including animals in the HCI community is not just beneficial for the animals. By studying and experimenting with HCI in a new context, one can gain insight and inspiration about how to adapt computers for use by biological entities. In the same way Gregg C. Vanderheiden encouraged designers to think of disabled and differently-abled users if for no other reason than the potential to increase understanding of the mainstream, thinking of interfaces in terms of animals could have similar benefits. For example, it was the construction of a device for dog that led us to speculate task-oriented instructional interactions might be an optimal mode of play between distant parents and their preverbal children.

We believe this thesis to be the first work to explicitly apply HCI to non-human animals. Although there has been previous research on human-animal communication, some of it mediated by a computer, none of this research has taken place under the auspices of human-computer interactions. By demonstrating the parallels between animal trainers and interface designers, we hope to bring awareness to both groups of people.

This thesis does more than present a theoretical framework for animal-computer interactions. We have built and tested a real world device, and addressed such non-theoretical issues as latency and bandwidth over public networks such as the Internet. The dogs and dog owners respond favorable to the constructed device, and have expressed interest in continuing use.

10 Appendix A: Technical Design

Rover@Home consists of three pieces of hardware and three pieces of software.

Hardware:

- Feeder
- Junction box
- Alley-Oop

Software

- Server software
- Client applet.
- Security access applet

10.1 Hardware

All hardware was initially prototyped on breadboards. Schematics and PC boards were designed with Protel 99SE. The PC board layout was sent to AP Circuits for fabrication. We used the AP Circuits P1 process that is less expensive, and turns boards around in about 2 days. Electronics were primarily purchased from Digikey.com and hardware primarily came from McMaster.com. Enclosures for the feeder and junction box were made of acrylic from Mr. Plastics in Sommerville, and cut in the laser cutter using CorelDraw 9.0 for design.

All three hardware components use Microchip Technology PIC chips. The Microchip MP Lab software from Microchip.com was used in conjunction with the CCM C-compiler to write the code in C and download it to the PIC chips. We used the Microsoft Visual Studio editor to develop the PIC chip C code for its customizability and context-sensitive coloring.

All three hardware devices employ in-circuit serial programming (ICSP). For the alley-oop and the feeder, the means of programming the PIC chip is through the same RJ-12 connector that is used to connect to the junction box. This reuse of the connector saves space, and allows reprogramming without disassembly. For the junction box, the ICSP connector is also RJ-12, but is internal and requires the removal of the top, which is fastened with two phillips-head screws.

The wiring diagram for the ICSP harness is depicted below, in Figure 10-1. Because the 16F873 used by the junction box and the 16F84 used by the alley-oop and feeder have different ICSP pins, two leads are necessary.

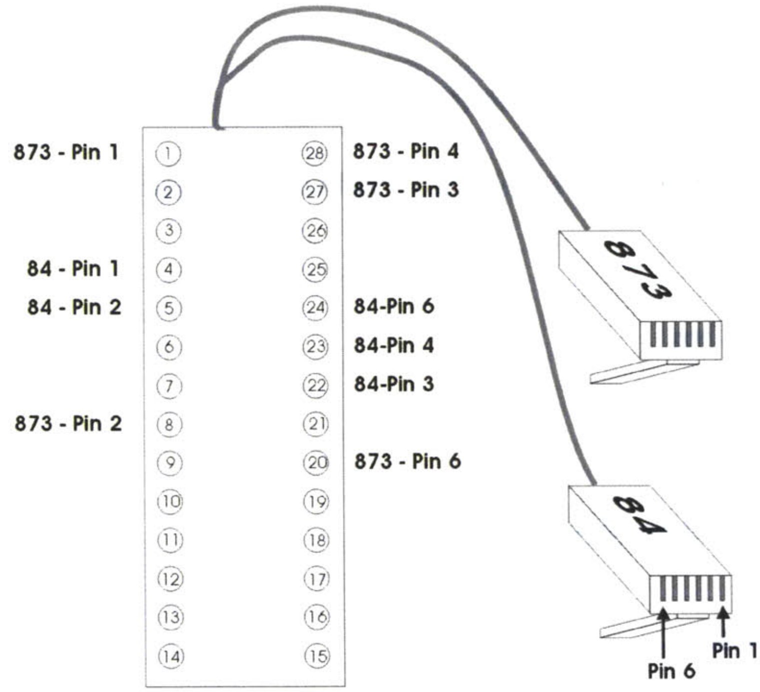


Figure 10-1: In-Circuit Serial Programming (ICSP) Wiring harness. The 28-pin DIP plugs into the Picstart Plus programmer, and the RJ-12 jacks plug into the circuit board. Because the 16F84 and 16F873 have different pin programming configurations, each gets its own cable. Note that pin 5 of the RJ-12 jacks are not used.

10.1.1 Junction Box

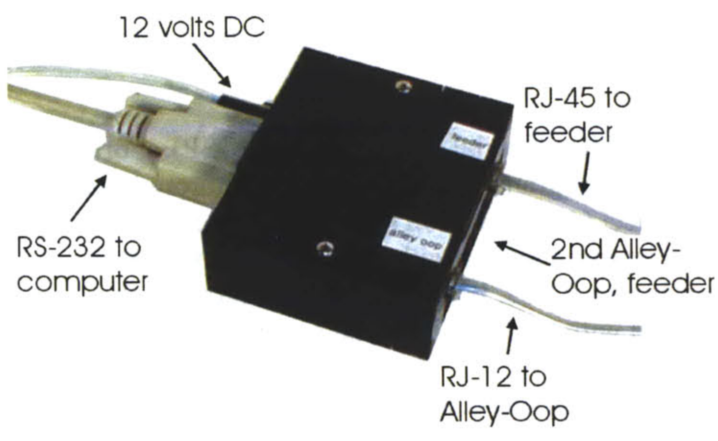


Figure 10-2: Junction box between the dog's computer and the Rover@Home peripherals. All cables are low voltage and present no danger to the dog. The cables to the Rover@Home peripherals are standard crossover telephone RJ-12 cables, easily replaceable from a wide range of vendors such as Radio Shack.

The junction box connects to the host PC over the serial port and is the glue between the host software and the Rover@Home peripherals. Rover@Home peripherals are attached to the junction box with standard 4-conductor RJ-12 telephone cables. Note these cables are standard "crossover" cables, not less common "straight" cables. A diode protects the peripherals from reverse-polarity damage so a non-crossover cable will not cause harm. However, a non-crossover cable will not allow peripherals to function, and there is currently no feedback to the user this is the cause of malfunction.

The junction box uses a more expensive 16F873 chip primarily because it includes a hardware UART for the RS-232 connection. This means if serial data comes in while the chip is busy with something else, the data does not get lost. Prototypes of the junction box with a 16F84, which lack the hardware UART, dropped data. The 16F873 also has three timers instead of one, analog inputs, and more memory, making this chip far more versatile for future applications. The 16F873 can be upgraded to the 16F876 which doubles the memory.

Each of the four RJ-12 jacks for attaching external peripherals contain four conductors:

- **+ 12 volts from wall transformer**. Note this requires the use of a 5-volt regulator on each peripheral. We decided to use 12 volts instead of 5 volts because peripherals such as the feeder require 12 volts for the motor, and this higher voltage gives peripherals more flexibility in conditioning the power after traveling over a potentially long cable. The low cost and small size of a typical 5-volt regulator such as a 7805 makes this a practical decision.

- **Ground**

- **Input line**. This takes input from the peripheral. It is tied to ground through a 10k resistor. This maintains the input in a low state in the event of no peripheral attached. Without this resistor, the junction box mistakes voltage fluctuations as signals from a peripheral.

- **Output line.** Write data to the peripheral

Restricting ourselves to four conductors when six are available with the RJ-12 jacks is important for two reasons:

- Connecting cable is smaller, more flexible, and more widely available. 6-conductor cables are bulkier than 4-conductor cables.

- For peripherals, RJ-12 jack needs to accommodate power and I/O lines, as well as in-circuit serial programming. ICSP requires five conductors, one of which is ground, and two of which can also be the I/O lines above.

Currently, the junction box will support two alley-oops and two feeders. Because the junction box merely passes commands between a particular peripheral and the computer, a wide range of devices can potentially be attached to the junction box. All that is required of peripherals is they operate from +12 volts, and communicate through a single input and output line.

Each of the four RJ-12 jacks for attaching external peripherals has its input and output lines wired to a dedicated pair of pins on the PIC chip. Peripherals do not share a common bus architecture. Additional peripherals cannot be added with a splitter. This greatly simplifies communication and eliminates the need for developing a databus sharing protocol. It also allows the junction box to identify what is attached to it merely by checking if something is attached to a particular jack.

The junction box can send out a "ping" to each of the four inputs and listen for a response. With this mechanism, the junction box can do automatic discovery of peripherals added and removed while the system is running. The junction box can also request each peripheral to send its version number.

The hardware was designed for fault tolerance. The junction box can be power cycled, and peripherals added and removed without restarting the software. Although not yet programmed, the hardware can support the software displaying the state and presence of each peripheral.

10.1.2 Feeder

Figure 10-3: Closeup of the custom designed feeder. The top is a 4" diameter x 12" high acrylic canister. The bottom is sawed off and attached to the base with four sheetmetal screws. The screws protrude into the canister and stir the food as the rotator turns, preventing food from simply turning with the rotator.

We initially used a commercial laboratory feeder from Med Associates. The device was rather large (12" x 9 ¼" x 7 ½), expensive (retails \$795 + \$125 for power supply), and made a noise that startled our first test dog. Therefore considerable resources were investing in developing our own custom pet food dispenser. The constraints upon us were:

- Low cost
- Small size
- Quiet
- Easy to clean

Factors that were not particularly important to us that may have been important to the Med Associates feeder are:

- Dose accuracy
- Dose latency (time between requesting food release and actual food release)

Inaccurate doses can paradoxically lead to increased display of behaviors. The fastest means to train an animal to exhibit a behavior is with a variable reinforcement schedule. Consistent reward levels for a behavior will inadvertently train an animal to do the minimum necessary to elicit reward. In contrast, a random reinforcement schedule will keep an animal uncertain about the threshold necessary to elicit a reward, and therefore make its display of the behavior more pronounced (Gray, 1994). So the feeder "bug" is actually a positive feature. The fact the dispensers emit from one to four pieces of food can be viewed as an intentional automatic variable reward. If the feeders were 100% reliable about dosage, we would probably be adding in a software feature to randomly vary the reward size.

Dose latency is also not a tremendous issue because the click is the reinforcer, not the food. Trick training depends of precise delivery of the click, not the food. As long as the food is within one or two seconds of the click, the animal's association between click and treat is maintained.

The feeder we designed is similar to the Med Associates. It is also similar to an automated gumball machine. A gearmotor turns a rotor until food is released through the PVC dispensing nozzle. An IR LED / detector pair inserted in the nozzle stops the motor once a treat has fallen through. The motor is timed to continue rotating a predetermined time so the openings of the rotor do not align with the nozzle. This prevents food from falling out when transporting or shaking the feeder. Dose size can be adjusted by changing the diameter of the holes in the rotor.

Food is held in a 4" acrylic food storage canister with the bottom sawed off. The mounting screws for the food holder are intentionally ¼" too long. By protruding into the food, they act as a stirrer, ensuring the food does not simply turn with the rotor.

A PIC chip controls the feeder. It listens for a signal coming in over the RJ-12 connector, starts the motor, waits for food to fall out, and sends a confirmation signal back to the controller. If no food falls out for five seconds, it stops the motor and sends an error signal to the controller so the user can take appropriate action.

Food released is detected with an IR LED / photodiode pair mounted in the dispensing nozzle. The food optically interrupts the beam, signaling the feeder to stop.

Power for the PIC comes over the RJ-12 connector. There is no power switch on the feeder. Being plugged into the junction box means it is on. This minimizes the chance of forgetting to turn the power switch ON.

10.1.3 Alley-Oop

For the alley-oop we modified an existing alley-oop from Gary Wilkes at http://www.clickandtreat.com. Because of the difficulty of disassembling alley-oops once they are put together, we worked with Gary Wilkes to obtain the raw alley-oop parts.

The alley-oops need to input when they are being tilted, and output the light and sound used for luring. Originally we sensed tilt with an Analog Devices ADXL202 connected to a PIC16F876 chip from Microchip Technology. These solid-state accelerometers detect two degrees of freedom over an entire hemisphere, but unfortunately the combination of the ADXL202 + PIC16F876 was insufficiently sensitive to detect the very small displacements of a dog touching the ally-oop. It was unclear whether the circuit noise was coming from the onboard A/D detectors, the ADXL202, noisy power supply, or some other factor. Powering the device with a battery and keeping all leads short helped reduce noise, but it was still above our target detection threshold.

Tilt detection of an entire hemisphere is not necessary, and paying $40 for such a chip is overdesign -- even for the Media Lab. Instead of pursuing the ADXL202, we fabricated a custom tilt sensor consisting of an arced enclosure containing a ¼" ball-bearing that optically interrupts an IR LED / detector pair (in fact, we used the same LED/photodetector pair as the feeder). This is a two-state device that only reports presence or absence of tilt. It is highly sensitive, low-cost, and does exactly what we need. Changing the degree of curvature of the arc in which the ball-bearing rests changes the sensitivity of the device.

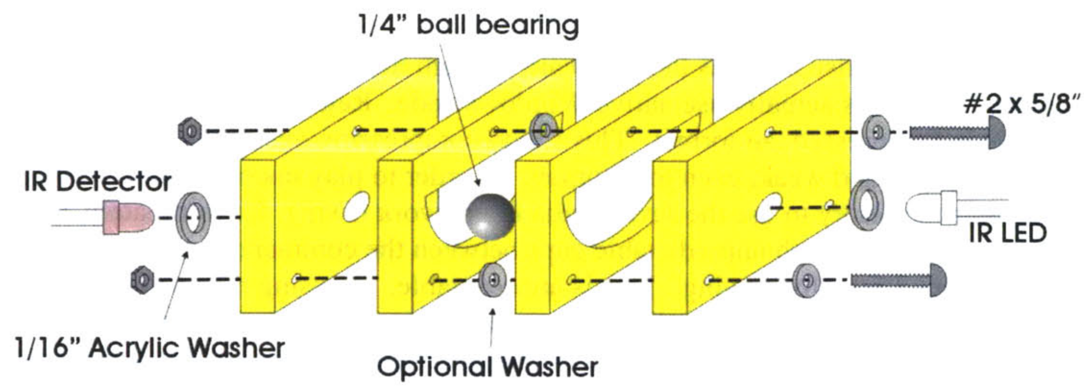


Figure 10-4: Exploded view of ball-tilt sensor. ¼ ball-bearing rolls back and forth in arc, and interrupts beam from IR LED to matching photodetector. Variations in thickness of the acrylic sometimes necessitate inserting a washer between the two inner layers of the housing. A single ¼" thick piece of acrylic could be used for the inner section instead of two 1/8" pieces, but beam spread in the lasercutter means cuts are not perfectly 90°. This causes the ball to fall to one side. Twoback-to-back 1/8" cutouts create a slight channel in which the ball can travel.

By using two of these ball-tilt sensors mounted at right angles, we can be assured that any displacement will be detected. It was suggested we use a single IR LED / detector pair mounted on the base of a bowl shaped enclosure. This would allow a single LED / detector pair to evaluate two dimensions. But this exposes us to missed displacements if the ball-bearing takes a path around the LED. Two 1D ball-tilt sensors at right angles ensure every displacement will be detected.

Similar to the feeder, alley-oop communication is controlled by a PIC chip. This chip sends signals to indicate X or Y tilt, and listens for signals to flash and / or beep. The alley-oop is powered by the connection with the junction box.

10.2 Software

The software consists of three parts, the server software that runs on the dog's computer, the Rover@Home applet that and runs on the client computer, and the security applet that restricts access to those without the correct password. The dog's computer serves both applets. All three were written in Java for its combination of ease and power. While cross-platform operation is certainly nice, it was not a primary consideration. In fact, we make use of some native code in the server software.

10.2.1 Server Software

The server code started life as a simple HTTP web server download from http://www.java.sun.com. The software listens for HTTP requests on port 80 (although users can change this port) and serves the appropriate page.

The server code also controls the hardware through the serial port, and releases the clicks. The clicks actually use native Windows code. Regular Java 1.3 only plays sounds in compressed .au format. This sound compression makes the click sound extremely dull and weak, even to a human. In order to play uncompressed wave files, it became necessary to use the Java Media Framework (JMF). Unfortunately, this proved unreliable, with unpredictable gaps between the command to play a sound, and the sound actually playing. This is unacceptable. By using native code, we can ensure prompt delivery of sounds. If we were to port to another platform, we could revert to JMF and hope it works better on that platform, or simply write native code for playing sound. The native code was developed in Microsoft Visual Studio 6.0.

The server has a log screen and a minimal UI for controlling the hardware. If the user chooses to operate Rover@Home locally, a connection through a local web-page works. Because the server supports multiple connections, forgetting to shut down the local web page will not prevent additional connections from remote sites. While only the first connection can receive NetMeeting video and audio, if the first connection is local, video is not transmitted, leaving the video connection available for future sessions.

10.2.2 Applet Software

When the client loads the Rover@Home web page, two connections back to the host computer are made. The first connection is through Microsoft NetMeeting embedded as an ActiveX control. The second connection is through a custom java applet that allows two-way communication with the server. This java applet acts as the interface for operating Rover@Home.

The java applet connects to the server over TCP/IP port 8081, therefore, this port must not be firewalled. Port 8081 is user-configurable in case this port is firewalled or in use by another application.

Microsoft NetMeeting connects over several TCP/IP ports, all of which must not be firewalled. Primary connection is made through ports (389, 522, 1503, 1720, 1731) as well as dynamically assigned UDP ports 1024-65535)[18]

The combination of ActiveX and Java 1.1 means we can only run on Internet Explorer 5.0 and later. Care was taken not to use Java beyond 1.1. Most notably, this meant avoiding the collection classes. Netscape Communicator 4.75 runs Java 1.1, but does not support the NetMeeting ActiveX control. Users content without video and audio can use Netscape without a problem. We anticipate future version of Netscape that can run NetMeeting will be able to run Rover@Home without trouble.

[18] Information taken from:
http://www.microsoft.com/Windows/NetMeeting/Corp/ResKit/Chapter4/default.asp. If this link no longer functions, try searching for "NetMeeting port firewall" from the Microsoft.com homepage.

When toggling checkboxes on the applet, the checkbox does not actually change state until confirmed by the server. This ensures feedback only happens after the server has acknowledged the mode change. For instance, when turning on "auto-trigger," the auto-trigger checkbox does not appear checked until the request has been sent to the server, and the server has returned an acknowledgement. During this interval, a red dot blinks in the lower left corner of the relevant button to give the user feedback the request is still pending.

This mechanism also ensures all copies of the webpage remain synchronized. Changing a state on one Rover@Home page changes the state for all web pages viewing a particular interaction.

All the compiled applet .CLASS files, as well as associated .GIF files, are packaged into a .JAR file for faster downloading. The .JAR files is 142 Kbytes.

10.2.3 Security

Rover@Home contains basic password protection encoding. When a user first connects to the server, they are served a logon applet that encrypts a username and password for validation by the server. A correct password validates that client IP address and allows downloading of the Rover@Home webpage. Leaving the Rover@Home webpage removes that IP address from the validation list, requiring the user to once again log on.

To prevent interception and repeat attacks, the server sends the logon applet a 20 byte random nonce, which is encoded, along with the username and password, with the SHA-1 message digest algorithm. This encoded string is sent to the server and challenged with a locally encoded copy. Because the nonce is generated fresh each time, repeat attacks will fail.

The password is stored on the user's computer after being encoded in the SHA-1 algorithm. Within the limits of decoding SHA-1 encoded messages, it is impossible for the user's password to be ascertained by a third party, even if they have access to the local computer.

In using NetMeeting, Rover@Home contains a huge security hole. It is possible to password protect NetMeeting, but the implementation is fundamentally flawed. When a user connects to a remote NetMeeting peer, the video stream is initiated, and the password dialog box appears over the video window. This allows the external party to view and hear the dog without entering a password.

Short of Microsoft fixing this problem, hacks involving native code could fill this security hole. NetMeeting is composed of COM objects that can be controlled by external applications. By placing the NetMeeting under control of the Rover@Home server, audio and video can be restricted until a suitable password has been entered.

11 Appendix B: Schematics and Lasercut

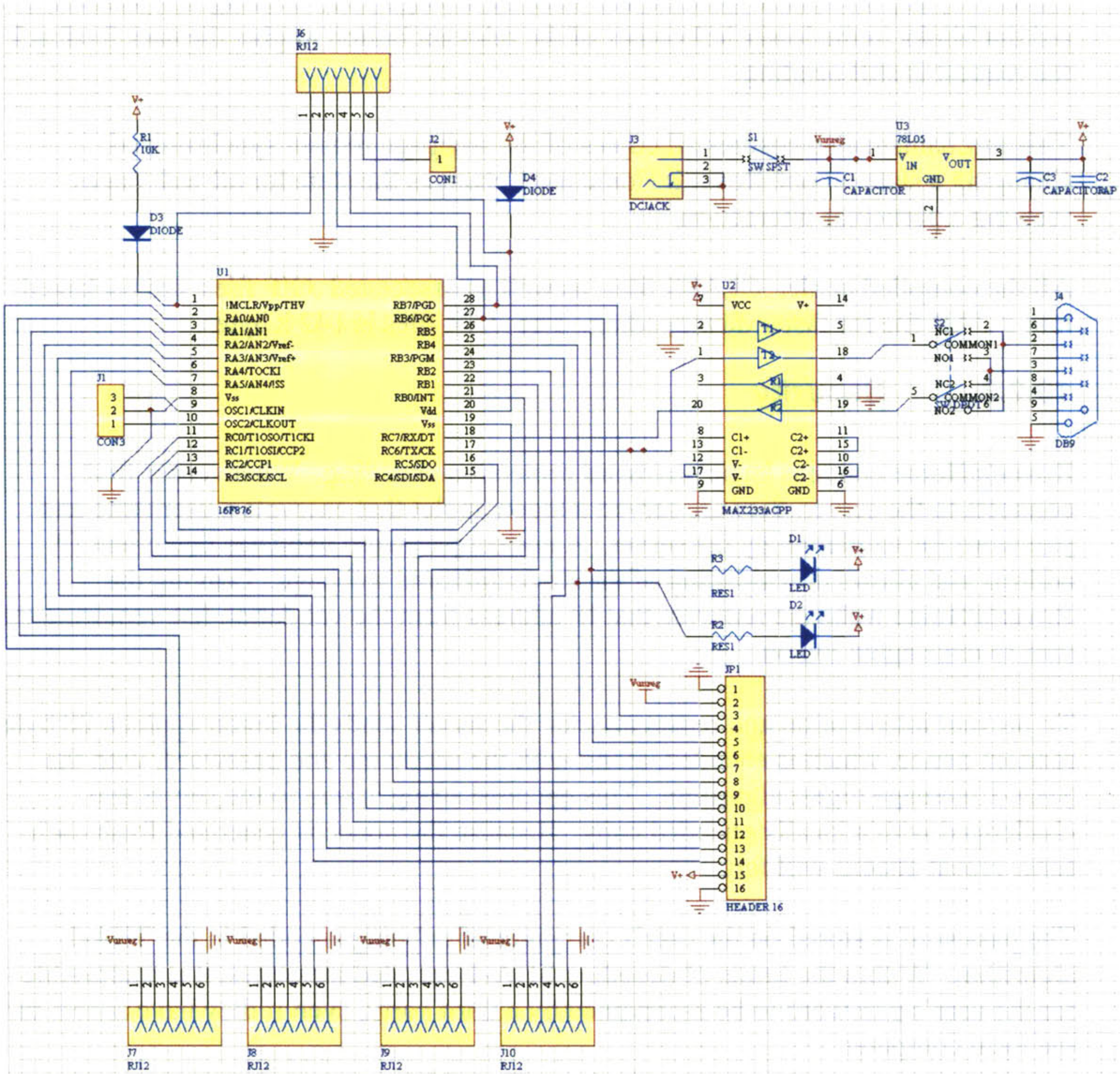


Figure 11-1: Junction Box. Note this schematic is missing 10k resistors between ground and pins A0, A2, B0, B2. These four resistors keep the peripheral lines low when nothing is plugged in. These four resistors were manually soldered on the board, and should be included in the next version of this board.

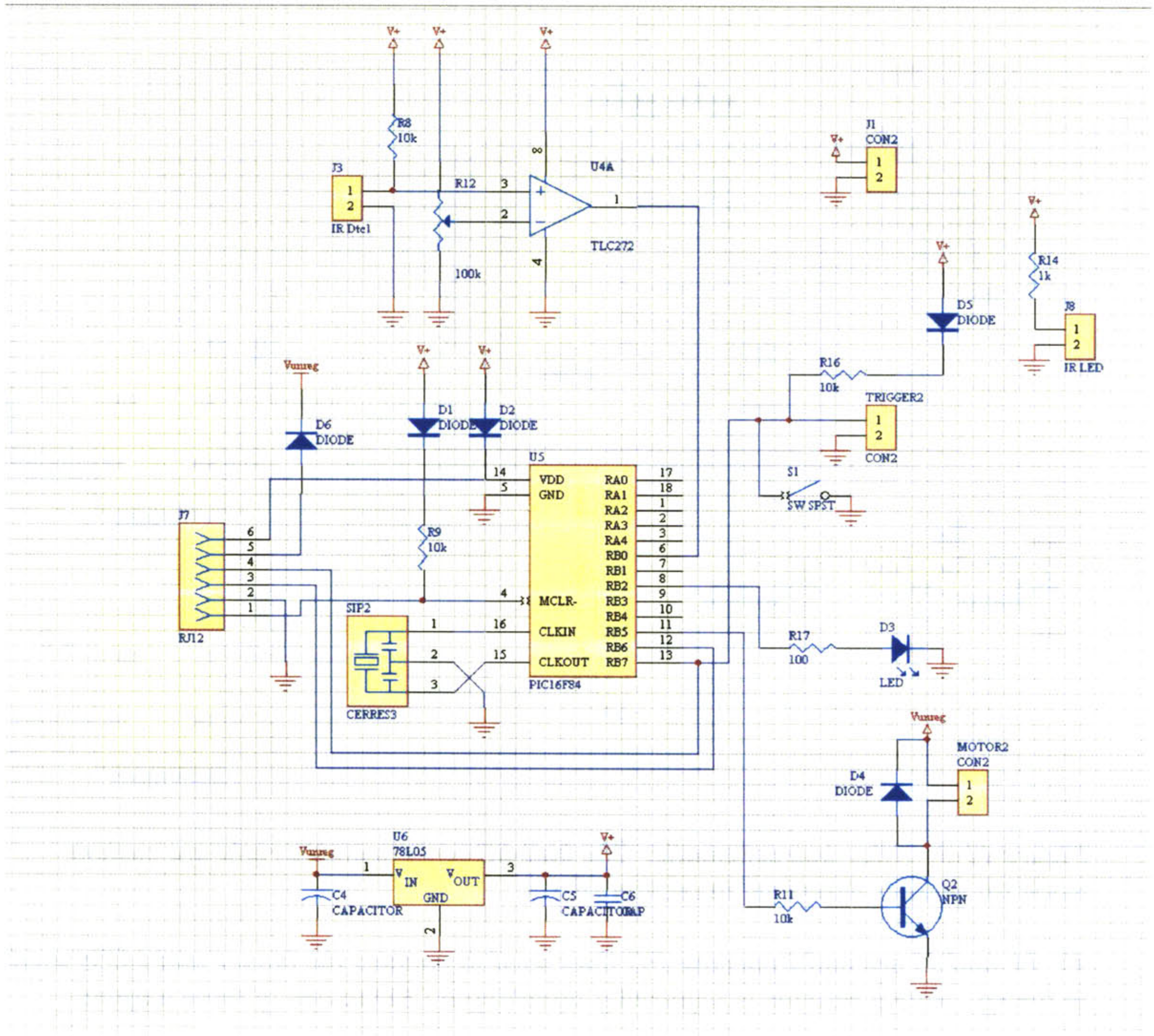


Figure 11-2: Feeder schematic. Note the RJ-12 jack on the left is used both to communicate with the junction box, and for the in-circuit serial programming (ICSP) connection with the Picstart Plus programmer.

The "trigger" connector is meant to manually start the feeder, but is attached to the input line from the junction box. It should be attached to an unused line of the PIC chip.

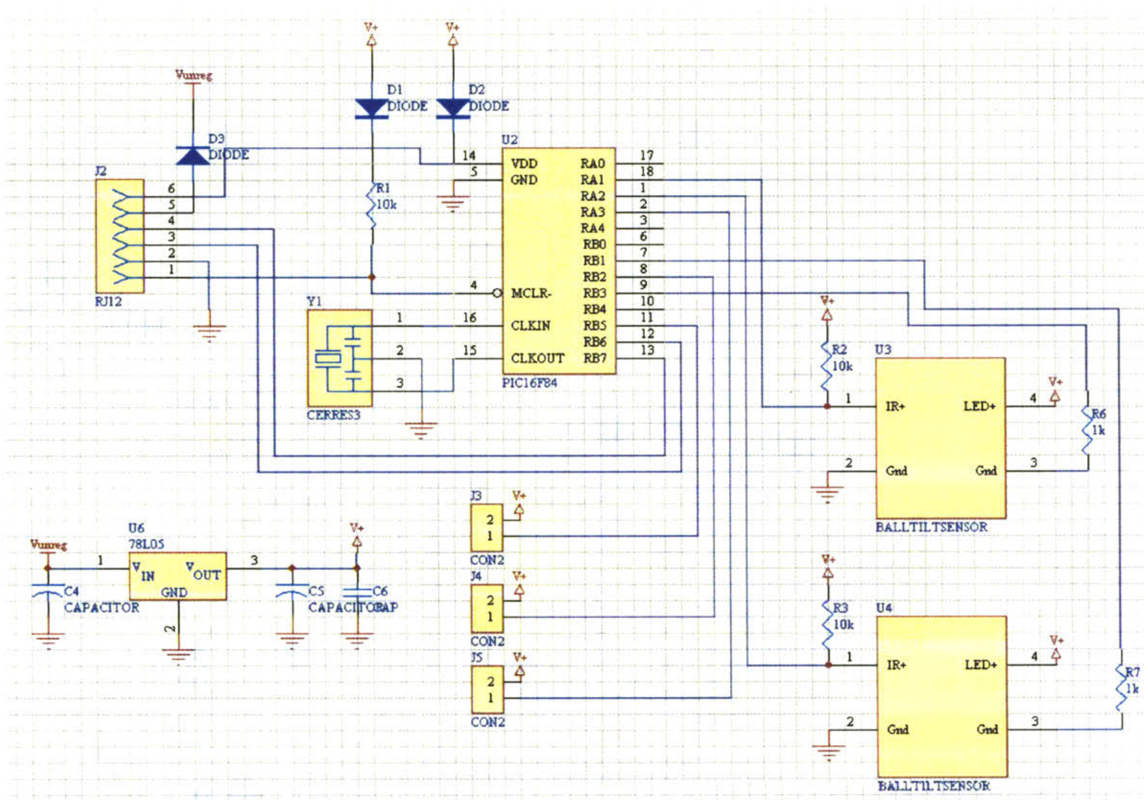


Figure 11-3: Alley-oop schematic. The ball-tilt sensors on the right are custom made tilt sensors described in section 10.1.3

J3, J4, and J5 are multi-purpose I/O lines used to connect the buzzer and LEDs used in luring. Note a resistor is required if attaching a LED to one of these outputs.

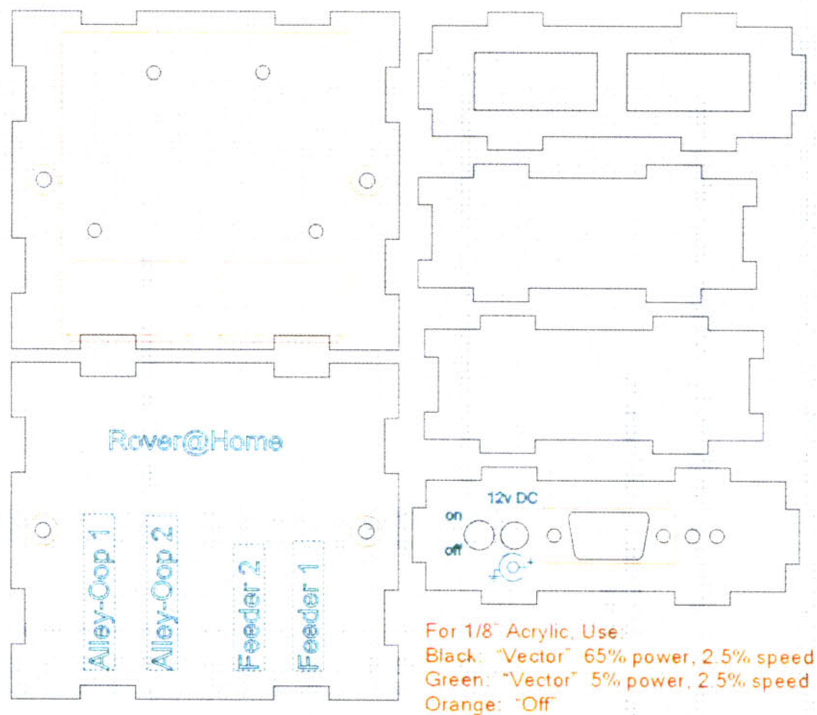


Figure 11-4: Lasercut of the junction box enclosure. Enclosure is made of 1/8" acrylic.

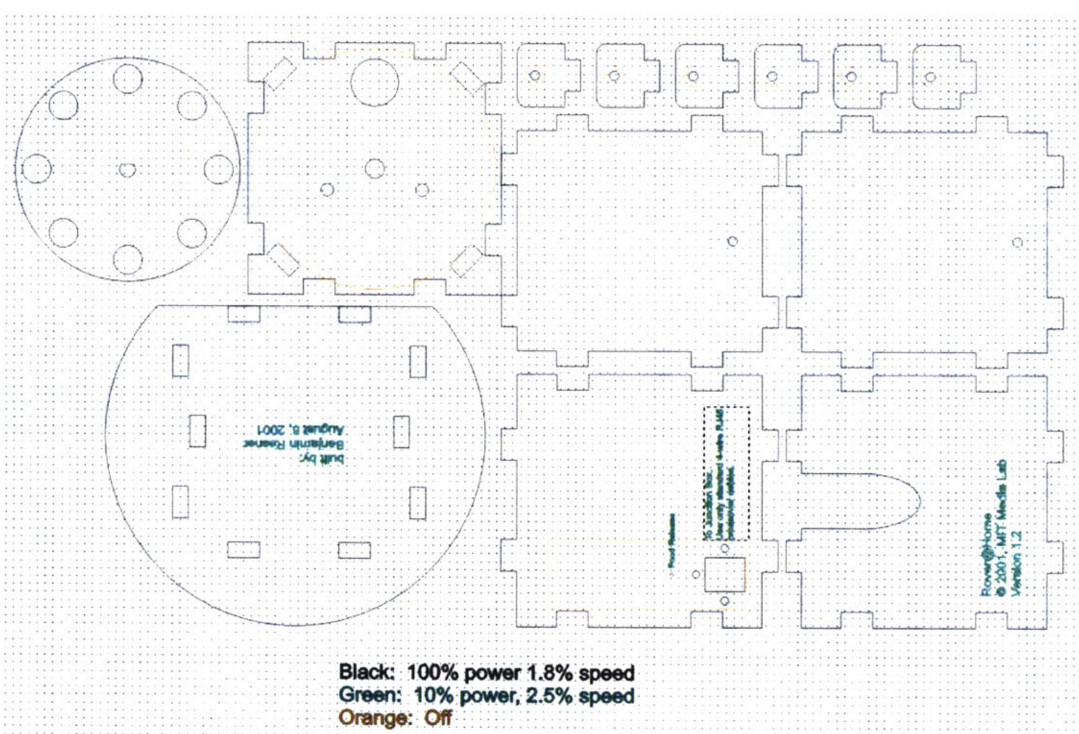


Figure 11-5: Lasercut of the feeder enclosure. Enclosure is made of ¼" acrylic.

12 Appendix C: Rover@Home Parts List

A complete Rover@Home installation, including the junction box, one alley-oop, one feeder, and all connecting cables costs about $180 when built in quantities of 10 – 25. Price will likely be significantly less in larger quantities. The largest single expense is $46.60 for the gearmotor for the feeder. This gearmotor used is more powerful than necessary, and can probably be sourced for much less. Other opportunities for cost savings include the use of less powerful microcontrollers (3 pics totaling $14.20) and less expensive housing (acrylic is easy to work with in the laser-cutter, but not cheap).

A complete prototype takes about 6 hours to assemble. Constructing prototypes in batches significantly reduces assembly time.

PART	SOURCE	PART NUMBER	QUANT	APPROX COST	EXTENDED COST	COMMENTS
ALLEY-OOP						
Alley-oop parts from Gary Wilkes	clickandtreat.com		1	$ 15.00	$ 15.00	
1.5" diameter clear acrylic hemispheres	FX Supply	VHH-48	2	$ 1.85	$ 3.70	http://www.fxsupply.com/products/hemis.html
Ball-Tilt Sensor			***2***			
1/8" acylic cutouts	Mr. Plastics	n/a	4	$ 0.10	$ 0.40	
1/16" acylic lasercut washers	Mr. Plastics	n/a	4	$ 0.10	$ 0.40	
1/4" ball bearing	Bocabearings.com		2	$ 0.10	$ 0.20	
#2 flat washers	McMaster.com	90945A705	2	$ 0.03	$ 0.06	
#2 lock washers	McMaster.com	92146A003	3	$ 0.02	$ 0.06	
#2-56 x 5/8 slotted pan-head screw	McMaster.com	91772A082	2	$ 0.03	$ 0.06	
#2-56 nut	McMaster.com	91841A003	2	$ 0.03	$ 0.06	
Electronics						
printed circuit board (base)	AP Circuits	Custom	1	$ 5.00	$ 5.00	Plus $50 batch setup cost
10 MHz 16F84 SOIC Package	Digikey.com	PIC16F84T-10/SO-ND	1	$ 4.80	$ 4.80	
10 MHz resonator	Digikey.com	X906-ND	1	$ 0.58	$ 0.58	
1N4148 Diode	Digikey.com	1N4148MSCT-ND	3	$ 0.05	$ 0.15	
7805 Regulator - TO92 package	Digikey.com	ZR78L05C-ND	1	$ 0.97	$ 0.97	
22 uf cap	Digikey.com	P814-ND	2	$ 0.25	$ 0.50	
.1 uf cap	Digikey.com	1210PHCT-ND	1	$ 0.10	$ 0.10	

10k	Digikey.com	10KQBK-ND	3	$	0.05	$	0.15	
1k	Digikey.com	1KQBK-ND	2	$	0.05	$	0.10	
IR LED, T-1 package 940 nm	Digikey.com	160-1028-ND	1	$	0.75	$	0.75	
IR Photodetector, T-1 package	Digikey.com	160-1030-ND	1	$	0.75	$	0.75	
Superbright LEDs	Photonmicrolight.com	Red LED	1	$	2.00	$	2.00	
Buzzer	Digikey.com	P9928-ND	1	$	6.75	$	6.75	
2-pin molex housing	Digikey.com	WM2000-ND	3	$	0.20	$	0.60	
2-pin molex connector	Digikey.com	WM4200-ND	3	$	0.16	$	0.48	
RJ-12 Side Entry Jack	Digikey.com	A9049-ND	1	$	0.57	$	0.57	
Double-sided foam tape	misc	n/a	1	$	0.20	$	0.20	
SIP plugs for ball-tilt sensor	Digikey.com	ED7064-ND	8	$	0.05	$	0.40	
TOTAL ALLEY-OOP						**$**	**44.79**	
JUNCTION BOX								
Fab								
1/8" acrylic	Mr. Plastics		1	$	2.00	$	2.00	$20 / sheet -- about 10 cases per sheet
3/16 nylon spacers for #4 screw	McMaster.com	94639A704	4	$	0.05	$	0.20	
#4-40 nut	McMaster.com	91395A005	4	$	0.03	$	0.12	Holds circuit board in case
#4-40 x 1" bolts	McMaster.com	91772A115	4	$	0.03	$	0.12	Holds circuit board in case
#4-40 x 1/2" bolts	McMaster.com	91772A110	4	$	0.03	$	0.12	Attaches covers to standoffs
#4-40 x 1" hex threaded standoffs	McMaster.com	90308A167	2	$	0.57	$	1.14	Attaches cover
#4-40 lock washer	McMaster.com	92146A530	2	$	0.03	$	0.06	
Rubber feet	Digikey.com	SR51-ND	4	$	0.20	$	0.80	
Electronics								
printed circuit board	AP Circuits		1	$	5.00	$	5.00	Plus $50 batch setup cost
10 MHz 16F873 DIP	Digikey.com	PIC16F873-20/SP-ND	1	$	4.60	$	4.60	
10 MHz resonator	Digikey.com	X906-ND	1	$	0.58	$	0.58	
1N4148 Diode	Digikey.com	1N4148MSCT-ND	2	$	0.05	$	0.10	
RJ-12 top entry	Digikey.com	A9084-ND	1	$	0.62	$	0.62	
RJ-12 side entry	Digikey.com	A9049-ND	4	$	0.57	$	2.28	
MAX233	Digikey.com	MAX233CPP	1	$	5.50	$	5.50	
DPDT switch	Digikey.com	SW102-ND	1	$	0.76	$	0.76	
DB9 Connector	Digikey.com	A2100-ND	1	$	2.11	$	2.11	
Red LED	Digikey.com	P363-ND	1	$	0.13	$	0.13	
Green LED	Digikey.com	P364-ND	1	$	0.13	$	0.13	
DC Power jack	Digikey.com	CP-002A-ND	1	$	0.25	$	0.25	

DC Power Plug	Digikey.com	CP-004A-ND	1	$	0.56	$	0.56	
SPDT power switch	Digikey.com	CKN1059-ND	1	$	3.60	$	3.60	
12 volt power supply, 500 ma	RadioShack.com	273-1773	1	$	12.99	$	12.99	
7805 regulator TO-220 package	Digikey.com	LM340T-5.0-ND	1	$	0.65	$	0.65	
22 uf cap	Digikey.com	P814-ND	2	$	0.13	$	0.26	
.1 uf cap	Digikey.com	1210PHCT-ND	1	$	0.09	$	0.09	
10k resistor	Digikey.com	10KQBK-ND	1	$	0.05	$	0.05	
100 ohm resistor	Digikey.com	100QBK-ND	2	$	0.05	$	0.10	
TOTAL JCN BOX						**$**	**44.92**	
FEEDER								
Fab								
1/4" black acrylic	Mr. Plastics	n/a	1	$	12.00	$	12.00	$36 / sheet -- about 3 feeders per sheet
1/8" acrylic (any color)	Mr. Plastics	n/a	1	$	0.50	$	0.50	
#8 x 5/8 mounting screws	McMaster.com	92470A196	6	$	0.03	$	0.18	
Gearmotor	McMaster.com	6331K35	1	$	46.60	$	46.60	Pittman GM9413-4 12VDC 6.8 RPM Gearmotor
#10-32 x 1" motor mounting screws	McMaster.com	91781A833	2	$	0.08	$	0.16	
4" diameter x 12" tall Acrylic Canister	US Acrylic	1850	1	$	4.00	$	4.00	http://www.usacrylic.com
#4-40 x 3/8 nylon spacers	McMaster.com	94639A710	2	$	0.03	$	0.06	
#4-40 x 1" bolts	McMaster.com	91772A115	2	$	0.03	$	0.06	
#4-40 nuts	McMaster.com	91395A005	2	$	0.03	$	0.06	
#4-40 lock washer	McMaster.com	92146A530	2	$	0.03	$	0.06	
Electronics								
printed circuit board	AP Circuits		1	$	5.00	$	5.00	Plus $50 batch setup cost
10 MHz 16F84 DIP PIC	Digikey.com	PIC16F84-10/P-ND	1	$	4.80	$	4.80	
10 MHz resonator	Digikey.com	X906-ND	1	$	0.58	$	0.58	
TLC272CP Op-Amp	Digikey.com	296-1824-5-ND	1	$	0.62	$	0.62	
RJ-12 top entry	Digikey.com	A9084-ND	1	$	0.62	$	0.62	
7805 regulator TO92 package	Digikey.com	ZR78L05C-ND	1	$	0.97	$	0.97	
22 uf cap	Digikey.com	P814-ND	2	$	0.65	$	1.30	
.1 uf cap	Digikey.com	1210PHCT-ND	1	$	0.09	$	0.09	
1N4148 Diode	Digikey.com	1N4148MSCT-ND	3	$	0.05	$	0.15	
1N5817 Schottky Diode	Digikey.com	1N5817CT	2	$	0.30	$	0.60	
Red LED	Digikey.com	P363-ND	1	$	0.13	$	0.13	
IR LED, T-1 package 940 nm	Digikey.com	160-1028-ND	1	$	0.22	$	0.22	
IR Photodetector, T-1 package	Digikey.com	160-1030-ND	1	$	0.33	$	0.33	

momentary "click" switch	Digikey.com	SW411-ND	1	$	0.27	$	0.27
2-pin Molex	Digikey.com	WM4200-ND	5	$	0.20	$	1.00
10k resistor	Digikey.com	10KQBK-ND	3	$	0.05	$	0.15
1k resistor	Digikey.com	1KQBK-ND	1	$	0.05	$	0.05
100 ohm resistor	Digikey.com	100QBK-ND	1	$	0.05	$	0.05
100k pot	Digikey.com	3386P-104-ND	1	$	1.12	$	1.12
TIP102 NPN Darlington Q (TO-220)	Mouser.com	511-TIP102	1	$	0.60	$	0.60
TOTAL FEEDER						**$**	**82.33**
MISC							
DB9 male - DB9 female serial cable	Digikey.com	AE1020-ND	1	$	4.69	$	4.69
RJ-12 6-pos, 4 wire crossover cable	Digikey.com	H1643-25-ND	2	$	3.00	$	6.00
TOTAL MISC						**$**	**10.69**
GRAND TOTAL						**$**	**182.73**

NOTES:

Costs are approximate for quantities around 10 - 25. Larger quantities can be significantly less.

5/8 sheetmetal screws are coated with food-grade lubricant -- McMaster part #1404K11

Acrylic is glued with Weld-On #16 high viscosity cement, McMaster.com part number 7515A11

Mr. Plastics is located at:
352 McGrath Hwy
Sommerville, MA
617-623-7000

13 Appendix D: Research Approval

This research has been approved under both the MIT Committee on Use of Humans as Experimental Subjects (COUHES), approval number 2765 on April 4, 2001, and the MIT Committee on Animal Care, approval number 01-020 on April 23, 2001.

14 Appendix E: Additional Information

Additional technical specifications, including an archival CD of all software and design can be obtained from the author – benres@media.mit.edu. This CD includes:

- Java files for server and applet, as well as supporting HTML, WAV, and GIF files.
- CPP and Microsoft Visual Studio project files for Windows native code.
- C files for PIC chips
- Additional build instructions
- Corel Draw lasercutter files for junction box and feeder enclosures
- Protel files and libraries for all circuit boards
- Text document detailing errors in circuit boards
- Excel spreadsheet of parts and approximate costs contained in Appendix C
- Source image files contained in this document
- Source Corel Draw images contained in this document
- Rover@Home AVI video
- Applications, Consent forms, and other verbiage from the MIT Committee on Animal Care (CAC) and MIT Committee on Use of Humans as Experimental Subjects (COUHES)
- This document in Microsoft Word 2000 format.
- Endnote bibliography file

15 Bibliography

American Pet Association (1998). Fun Pet Statistics, http://www.apapets.com/petstats2.htm

Apple Computer Inc. (1992). Macintosh Human Interface Guidelines, Addison-Wesley.

Arbitron Inc & Coleman Associates (2001). Broadband Revolution 2 - The Media World of Speedies.

Barker, S. and Barker, R. (1988). "The human-canine bond: Closer than family ties?" Journal of Mental Health Counseling **10**: 46-56.

Beaver, B. V. (1999). Canine Behavior: A Guide for Veterinarians, W.B. Saunders Company.

Beck, A. and Katcher, A. (1996). Between Pets & People: The Importance of Animal Companionship. West Lafayette, Indiana, Purdue University Press.

Benford, S., Greenhalgh, C., et al. (1998). "Understanding and Constructing Shared Spaces with Mixed-Reality Boundaries." ACM Transactions of Computer-Human Interaction **5**(3): 185-223.

Beyer, H. and Holtzblatt, K. (1997). Contextual Design: A Customer-Centered Approach to Systems Design, Morgan Kaufmann Publishers.

Böttcher, P. (1999). The Visible Animal Project, http://www.vetmed.uni-muenchen.de/anat1/VisibleAnimal/VAP.html

Brave, S. and Dahley, A. (1997). "inTOUCH: A Medium for Haptic Interpersonal Communication." Proceedings of CHI'97, ACM Press.

Brave, S., Ishii, H., et al. (1998). "Tangible Interfaces for Remote Collaboration and Communication." Proceedings of the CSCW '98.

Breland, K. and Breland, M. (1951). "A Field of Applied Animal Psychology." American Psychologist **6**: 202-204.

Breland, K. and Breland, M. (1961). "The Misbehavior of Organisms." American Psychologist **16**: 681-684.

Brown, P. L. (2001). The Warp and Woof of Identity Politics for Pets. The New York Times. New Your.

Budiansky, S. (2000). The Truth About Dogs. New York, Viking.

Burch, M. R. and Bailey, J. S. (1999). How Dogs Learn. New York, Howell Book House.

Buxton, W. (1995). "Integrating the Periphery and Context: A New Model of Telematics." Proceedings of Graphics Interface '95: 239-246.

Chang, A., Koerner, B., et al. (2001). "LumiTouch: An Emotional Communication Device." Proceedings of CHI'01, ACM Press.

Chen, D. and Haviland-Jones, J. (1999). "Rapid Mood Change and Human Odors." Physiology & Behavior **68**: 241-250.

Chen, D. and Haviland-Jones, J. (2000). "Human Olfactory Communication of Emotion." Perceptual and Motor Skills **91**.

Clark, M., Ferrara, T., et al. (1990). Koko's MAC II. The Art of Human-Computer Interface Design. B. Laurel, Addison Wesley: 95-102.

Clutton-Brock, J. (1995). Origins Of The Dog: Domestication and Early History. The Domestic Dog: It's Evolution, Behavior and Interactions with People. J. Serpell. Cambridge, Cambridge University Press: 7-20.

Communication Technology Lab at Michigan State University (1995). The Personal Communicator, http://commtechlab.msu.edu/products/pcom/

Consumer Reports (1998). Feeding Fido (and Fifi, Fluffy, Max...). Consumer Reports.

Copinger, R. and Schneider, R. (1995). Evolution of Working Dogs. The Domestic Dog: It's Evolution, Behavior and Interactions with People. J. Serpell. Cambridge, University of Cambridge Press: 21-47.

Coren, S. (2000). How To Speak Dog: Mastering the Art of Dog-Human Communication. New York, The Free Press, A division of Simon & Schuster, Inc.

Cyrano Sciences (2001). Cryanose 320 Specifications, http://cyranosciences.com/

Digisents.com (2001). iSmell, http://www.digiscent.com

Druin, A. (1999). "Cooperative Inquiry: Developing New Technologies for Children with Children." Proceedings of CHI'99, ACM Press.

Duin, J. (1998). Now it's cholesterol-free low-cal snacks...for dogs. Washington Times Corporation.

Egido, C. (1988). "Videoconferencing as a Technology to Support Group Work: A Review of its Failure." Association for Computing Machinery.

Farlow, S. (2001). Personal Communication. Lincoln.

Fish, R. S., Kraut, R. E., et al. (1992). "Evaluating Video as a Technology for Informal Communication." Communications of the Association for Computing Machinery.

Fitts, P. M. (1954). "The Information Capacity of the Human Motor System in Controlling the Aplitude of Movement." Journal of Experimental Psychology **47**: 381-391.

Fleishman, L. I., McClintock, W. I., et al. (1997). "Color Perception and the Use of Video Playback Experiments in Animal Behaviour." Animal Behaviour **56**: 1035-1040.

Fogle, B. (1990). The Dog's Mind: Understanding Your Dog's Behavior, Howell Book House.

Fouts, R. (1997). Next of Kin: My Conversations With Chimpanzees. New York, Avon Books, Inc.

Friedmann, E. and Thomas, S. (1995). "Pet ownership, social support, and one-year survival after acute myocardial infarction in the Cardiac Arrhythmia Suppression Trial (CAST)." American Journal of Cardiology **76**(17).

Gannett News Services (2001). Internet User Statistics.

Gardner, R. A., Gardner, B. T., et al., Eds. (1989). Teaching Sign Language to Chimpanzees. New York, State University of New York.

Gorilla Foundation (2001). Welcome to the Gorilla Foundation, http://www.koko.org/

Gray, P. (1994). Psychology, 2nd Edition. New York, Worth Publishers.

Grenier, R. (2000). The Difficulty of Being a Dog. Chicago, University of Chicago Press.

Iovine, J. V. (2000). At an Indoor Pet Spa, Why Let the Dogs Out? The New York Times. New York.

Isaacs, E. A. and Tang, J. C. (1993). "What Video Can and Can't Do For Collaboration: A Case Study." Proceedings of the first ACM International Conference on Multimedia.: 199-206.

Ishii, H. and Kobayashi, M. (1993). "Integration of Interpersonal Space and Shared Workspace: ClearBoard Design and Experiments." Proceedings of CHI'93, ACM Press.

Kaye, J. (1999). Symbolic Olfactory Display. Media Arts and Sciences. Cambridge, MIT: 144.

Kinetic Strategies (2001). Broadband Internet Subscriber Base Tops 9 Million. Phoenix, Kinetic Strategies.

Kuhn, T. S. (1962). The Structure of Scientific Revolution. Chicago, University of Chicago Press.

Lynn, S. (2001). Personal Communication.

Mackay, W., Svendsen, E. H., et al. (2001). "Who's In Control? Exploring Human-Agent Interaction in the McPie Interactive Theater Project." Conference in Human Factors in Computing Systems.

McNicholas, J. and Collis, G. M. (2000). "Dogs as Catalysts For Social Interaction: Robustness of the Effect." Brittish Journal of Psychology **91**(Pt1): 61-70.

Miklosi, A., Polgardi, R., et al. (2000). "Inentional Behavior in Dog-Human Communication: An Experimental Analysis of "Showing" Behavior in the Dog." Animal Congnition **3**: 159-166.

Norman, D. and Draper, S. W. (1986). User Centered System Design. New Jersey, Lawrence Erlbaum Associates.

Nussbaum, D. (1999). Communication Choices with Deaf and Hard of Hearing Children, http://clerccenter.gallaudet.edu/SupportServices/series/4010.html

O'Connell-Rodwell, C. E., Arnason, B. T., et al. (2001). "Seismic properties of Asian elephant (Elephas maximus) vocalizations and locomotion." Journal of the Acoustical Society of America **108**(3066).

Pearsall, M. D. and Verbruggen, H. (1982). Scent: Training to Track, Search, and Rescue. Loveland, CO, Alpine Publications.

Pederson, E. R. and Sokoler, T. (1997). "AROMA: Abstract Representations of Presence Supporting Mutual Awareness." Proceedings of CHI'97, ACM Press.

Pepperberg, I. M. (1999). The Alex Studies: Cognitive and Communicative Abilities of African Grey Parrots. Cambridge, Harvard University Press.

Porter, R. H. (1999). "Olfaction and Human Kin Recognition." Genetica **104**: 259-263.

Preece, J. (1994). Human-Computer Interaction. Essex, Addison-Wesley.

Pryor, K. Clicker Magic (Video). Waltham, Sunshine Books, Inc.

Pryor, K. (1984). Don't Shoot The Dog: The New Art of Teaching and Training, Bantam Doubleday.

Ramirez, K. (1999). Animal Training. Chicago, Shedd Aquarium.

Reiss, D. and McCowan, B. (1993). "Sponteneous Vocal Mimicry and Production by Bottlenose Dolphins (Tursiops truncatus): Evidence for Vocal Learning." Journal of Comparative Psychology **107**(3): 301-312.

Resner, B. (2001). Personal Communication.

Rine, J. (2001). Dog Genome Project, http://mendel.berkeley.edu/dog.html

Roony, N. J., Bradshaw, J. W., et al. (2001). "Do Dogs Respond to Play Signals GIven By Humans?" Animal Behaviour **61**: 715-277.

Rumbaugh, D. M., Ed. (1977). Language Learning by a Chimpanzee: The Lana Project. London, Academic Press.

Sanders, C. (1999). Understanding Dogs -- Living and Working with Canine Companions. Philadelphia, Temple University Press.

Savage-Rumbaugh, S. and Lewin, R. (1994). Kanzi: The Ape at the Brink of the Human Mind. New York, John Wiley & Sons, Inc.

Scott, J. P. and Fuller, J. L. (1965). Genetics and the Social Behavior of the Dog: The Classic Study. Chicago, University of Chicago Press.

Shettleworth, S. J. (1998). Cognition, Evolution, and Behavior, Oxford University Press.

Smith, R. S., Doty, R. L., et al. (1993). "Smell and Taste Function in the Visually-Impaired." Perception & Psychophysics **54**(5): 649-655.

Spector, M. (1999). Clicker Training for Obedience. Waltham, Sunshine Books.

Strong, R. and Gaver, B. (1996). "Feather, Scent, and Shaker: Support Simple Intimacy." Proceedings of CSCW '96.

Sweet, M. (1999). Sound Values: Check Frequency Response, Power & Dolby Compatiblity. SmartComputing. **7**.

Syrotuck, W. G. (1972). Scent and the Scenting Dog. Canastota, NY, Arner Publications.

Tollmar, K., Junestrand, S., et al. (2000). "Virtually Living Together: A Design Framework for New Communication Media." Conference proceedings on Designing interactive systems: processes, practices, methods, and techniques.

Tribune Business News (2000). Phoenix-Based Online Pet Supplies Retailer to Begin Selling Services. Tribune Business News.

Ullmer, B. and Ishii, H. (2000). "Emerging Frameworks for Tangible User Interfaces." IBM Systems Journal **39**(3&4).

Vanderheiden, G. C. (2001). Why Do We? Why Can't We? Future Perspectives and Research Directions. Closing Plenary Address of the SIG-CHI 2001 Conference, Thursday April 5, 2001.

Weiser, M. (1991). "The Computer for the 21st Century." Scientific American.

Wells, D. L. and Hepper, P. G. (2000). "The Discrimination of Dog Odours by Humans." Perception **29**: 111-115.

Wilkes, G. (1995). Click! & Treat Training Kit, Version 2.0, Gary Wilkes.

Wilkes, G. (2001). Personal Communication.

Williams, E. (2001). The Pet Loss Grief Support Website, http://www.petloss.com

Zak, O. (1999). Methods of Communication with the Deaf, http://www.zak.co.il/deaf-info/old/methods.html#list