Chapter 9
Opportunities and Challenges for Augmented Environments: A Distributed Cognition Perspective

James D. Hollan and Edwin L. Hutchins

Abstract Currently a new generation of inexpensive digital recording devices and storage facilities is revolutionizing data collection in behavioral science, extending it into situations that have not been typically accessible and enabling the examination of the fine details of action captured in meaningful settings. The ability to record and share such data has not only created a critical moment in the practice and scope of behavioral research but also presents unprecedented opportunities and challenges for the design of future augmented environments. In this chapter, we discuss these challenges and argue that fully capitalizing on the associated opportunities requires theoretical and methodological frameworks to effectively analyze data that capture the richness of real-world human activity. We sample five recent research projects from our laboratory chosen to exemplify a distributed cognition perspective and highlight opportunities and challenges relevant to the design and evaluation of augmented environments.

9.1 Introduction

There is currently a shift in cognitive science toward a view of cognition as a property of systems that are larger than isolated individuals. This extends the reach of cognition to encompass interactions between people as well as interactions with resources in the environment. As a consequence, the human body and material world take on central rather than peripheral roles. As Andy Clark put it, "To thus take the body and the world seriously is to invite an emergentist perspective on many key phenomena – to see adaptive success as inhering as much in complex interactions among body, world, and brain as in the inner processes bounded by the skin and skull." (Clark 1997) This new perspective is emerging from the fields of

J.D. Hollan (✉)
Department of Cognitive Science, Distributed Cognition and Human-Computer Interaction Lab, University of California, San Diego, CA, USA
e-mail: hollan@cogsci.ucsd.edu

S. Lahlou (ed.), *Designing User Friendly Augmented Work Environments: From Meeting Rooms to Digital Collaborative Spaces,* Computer Supported Cooperative Work,
DOI 10.1007/978-1-84800-098-8_9,

distributed cognition (Hutchins 1995a, b; Goodwin 2000; Hollan et al. 2000), embodied interaction (Clark 1997; Nunez 1996; Dourish 2001), and dynamical cognition (Port and van Gelder 1995; Thelen and Smith 1994).

Our research group at UCSD is dedicated to developing the theoretical and methodological foundations engendered by adopting this broader view of cognition and interaction and understanding how it can support the design of effective computer-mediated environments. Research in our lab ranges across cognitive science. We are particularly interested in understanding interactions among people and technology. Our work combines ethnographic observation and controlled experimentation to support theoretically informed design of digital work materials and collaborative work environments. Members of our lab are united in the belief that distributed cognition is a particularly fertile framework for designing and evaluating augmented environments and digital artifacts. A central image for us is workplaces in which people pursue their goals in collaboration with the elements of the social and material world. Our core research efforts are directed at understanding such environments: what we really do in them, how we coordinate our activity in them, and what role technology should play in them.

Currently a new generation of inexpensive digital recording devices and storage facilities is revolutionizing data collection in behavioral science, extending it into situations that have not been typically accessible and enabling the examination of the fine details of action captured in meaningful settings. Researchers from many disciplines are beginning to take advantage of increasingly inexpensive digital video and storage facilities to assemble extensive data collections of human activity captured in real-world settings. The ability to record and share such data has not only created a critical moment in the practice and scope of behavioral research but also presents unprecedented challenges and opportunities for the design of future augmented environments.

In this chapter, we discuss these challenges and argue that to fully capitalize on the associated opportunities three main issues must be addressed: (1) developing the theoretical and methodological frameworks required to effectively analyze rich real-world behavioral data, (2) reducing the huge time investments currently required for analysis and (3) understanding how to visualize and coordinate analyses focused at different scales so as to profit fully from the complementary perspectives of multiple disciplines. We first discuss the underlying behavioral challenge and then briefly introduce distributed cognition, the theoretical foundation for our work, and cognitive ethnography, the methodological approach we are developing. We then focus on examples (projects) from our lab to exemplify the approaches we are taking for capturing and analyzing data from real-world settings. We end the chapter with a set of lessons for the design of augmented environments.

9.2 A Challenge for Behavioral Science

What conditions can facilitate rapid advances and breakthroughs in behavioral science to rival those seen in the biological and physical sciences in the past century?

The emergence of cognitive science and the converging view across multiple disciplines that human behavior is a complex dynamic interaction among biological, cognitive, linguistic, social, and cultural processes are important first steps. While empirical and theoretical work is rapidly advancing at the biological end of this continuum, understanding such a complex system also necessitates data that capture the richness of the real-world human activity and analytic frameworks that can exploit that richness. This is important because to understand the dynamics of human and social activity, we must first understand the full context of those activities and this can only be accomplished by recording and analyzing data of real-world behavior. While such data are certainly needed, mere data cannot be the whole answer, since many researchers already feel that they are drowning in data. Data without appropriate theoretical and analytical frameworks do not lead to scientific advances.

Fortunately the revolution in digital technology can be coupled with exciting recent developments in cognitive theory. While these developments also heighten the importance of understanding the nature of real-world activities, they are in addition beginning to provide an analytic framework for understanding how cognition is embedded in concrete contexts of human activity. As described earlier, cognition is increasingly viewed as a process that extends beyond the skin and skull of the individual (Pea 1993; Hutchins 1995a, b; Cole 1996; Nardi 1996; Hollan et al. 2000; Clark 2003; Rogoff 2003). This shift in framing the unit of analysis for cognition introduces a host of previously overlooked cognitive phenomena to be documented, studied and understood.

Rich new digital data sources coupled with this shift in theory promise to advance the understanding of the links between what is in the mind, and what the mind is in. For example, just as widespread availability of audio tape recording supported the development of conversational analysis (Goodwin and Heritage 1990; Hutchby and Wooffitt 1998; Prevignano and Thibault 2003) and the ethno graphy of speaking (Gumperz and Hymes 1986; Bauman and Sherzer 1989), the advent of inexpensive digital video is starting to have a fundamental impact on cognitive science. The ability to record, view, and re-view the fine details of action in meaningful settings has made it possible to examine the phenomena at the core of embodied (Varela et al. 1991; Brooks 1991; Thelen and Smith 1994; Clark 1997; Lakoff and Johnson 1999), situation (Suchman 1987; Lave 1988; Brown et al. 1989; Chaiklin and Lave 1996; Clancy 1997) and distributed cognition (Hutchins 1995a, b; Clark 2001). The rise of studies in gesture in the past decade was also made possible by these technological changes and it is now transforming fields such as linguistics (McNeill 2005) and education (Goldin-Meadow 2003).

Sensor technologies starting to be deployed in augmented environments and new computational algorithms promise to further extend this transformation by enabling automatic recognition, tracking, and summarization of the meaningful components of the audio-video data (Zhao et al. 2003; Jones and Jones 2004). Thus, changes in theory give us new phenomena to see and provide new relevance to things already seen. Developments in digital technology create potential for new tools with which to see those things (Hollan et al. 1997; Card et al. 1999). Understanding these changes and developments are central to advances in behavioral science and these advances in turn provide a scientific foundation for the design of augmented environments.

9.3 Distributed Cognition

Distributed cognition promises to be a fertile theoretical framework for exploring the dynamics of human activity and for providing a foundation for design of augmented environments. It extends conceptions of cognitive processes beyond individual interaction and challenges key implicit presuppositions of current views. Like any cognitive theory, distributed cognition seeks to understand the organization of cognitive systems. And like most of cognitive science, it takes cognitive processes to be those that are involved in memory, decision making, inference, reasoning, learning, and so on. What distinguishes distributed cognition from other approaches is the commitment to two related theoretical principles.

The first of these principles concerns the boundaries of the unit of analysis for cognition. In every area of science, the choices made concerning the boundaries of the unit of analysis have important implications. Boundaries are often a matter of tradition in a field. In distributed cognition, one expects to find a system of systems that can dynamically configure itself to bring sub-systems into coordination to accomplish various functions. A cognitive process is delimited by the functional relationships among the elements that participate in it, rather than by the spatial co-location of the elements. Sometimes the traditionally assumed boundaries of the individual are exactly right. For other phenomena, however, these boundaries are not right because they either span too much or too little. Distributed cognition looks for cognitive processes wherever they may occur and does that looking on the basis of the functional relationships of elements that participate together in the process. A process is not cognitive simply because it happens in a brain, nor is a process non-cognitive simply because it happens in the interactions among many brains. For example, we have found it productive to consider small socio-technical systems such as the bridge of a ship (Hutchins 1995a, b), an airline cockpit (Hutchins 1995a, b; Hutchins and Klausen 1996), and as we discuss later, an automobile, a law office, and an augmented environment to assist interaction between deaf patients and physicians, as the unit of analysis.

The second principle concerns the range of mechanisms that may be assumed to participate in the cognitive processes. Whereas psychology looks for cognitive events in neural events inside individual actors, distributed cognition looks for a broader class of cognitive events and does not expect all such events to be encompassed by the skin or skull of an individual. For example, an examination of memory processes in an airline cockpit shows that memory involves a rich interaction between internal processes, the manipulation of objects, and the traffic in representations among the pilots. A complete theory of individual memory by itself is insufficient to understand how this memory system works. And a complete theory of internal cognitive functioning by itself is insufficient to understand the dynamics of human activity in augmented environments.

9.3.1 Distributed Cognition and Human Activity

When one applies these principles to the observation of human activity *in the wild*, at least three interesting kinds of distribution of cognitive process become apparent: (1) Cognitive processes may be distributed across the members of a social group. Tracking these processes produces insights about the dynamics of the social processes. (2) Cognitive processes may be distributed in the sense that the operation of the cognitive system involves coordination between internal and external (material or environmental) structure. Tracking these processes produces insights about the dynamics of agent/environment relations and is particularly relevant for understanding and designing augmented environments. (3) Processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events. Tracking these processes produces insights about the dynamics of social and cultural systems on longer timescales. The effects of these distributions of process are extremely important to an understanding of human cognitive accomplishments as products of human social dynamics and as a basis for building effective augmented environments.

9.3.2 Embodied Cognition

Distributed cognition theory embraces the movement in cognitive science toward a conception of embodied cognition. From this perspective, the organization of mind – both in development and in operation – is an emergent property of interactions among internal and external resources. In this view, the human body and the material world take on central rather than peripheral roles. For the understanding of human social dynamics, this means that elements of the social and material environment are more than simply stimuli for a disembodied cognitive system. Social and material patterns become elements of the cognitive system itself. This theoretical perspective promises an intellectual basis for a paradigm shift in thinking about human social dynamics and the design of augmented environments; one that takes material and social structures to be elements of cognitive systems and views on-going activity as a continually renegotiated emergent product of interaction.

9.3.3 Cognition and Culture

With the much more intimate relation between mind and environment that is provided by distributed cognition theory, comes the possibility of seeing new kinds of relations between culture and cognition. Hutchins treats this at length in his book, *Cognition in the Wild* (Hutchins 1995a, b). These new relations appear when we address the functional specifications for human cognition. What is a mind really used for? How are thinking tasks really done in the everyday world?

Permitting the boundary of the unit of analysis to move out beyond the skin situates the individual as an element in a complex cultural environment. In doing this, we find that cognition is no longer isolated from culture or separate from it. Where cognitive science traditionally views culture as a body of content on which the cognitive processes of individual persons operate, in the distributed cognition perspective, culture, in the form of a history of material artifacts and social practices, shapes the dynamics of the cognitive systems that transcend the boundaries of individual persons.

A central idea in distributed cognition is the notion of intelligence as an emergent property of interactions. This idea is reinforced by the connectionist challenge to traditional models of cognitive processing, and we have developed connectionist models of the emergence of structure in the interactions among networks in a community of networks. Connectionism, however, says nothing about the marginalization of the body and world. So to the idea of emergence, we add the idea that persons are embedded in complex environments that can be seen as active resources for learning, problem solving and reasoning. Culture is a process that accumulates partial solutions to frequently encountered problems. We live with the residue of previous activity and that is both enabling and constraining. Both culture and biology work this way and this fundamental fact gives life and mind a dynamic signature that is not seen in strictly physical systems (Thompson 2007). The intellectual tools that culture provides enable us to accomplish things that we could not do without them. At the same time, though they may blind us to other ways of thinking and make some things seem impossible, culture is a process that involves the interactions of mental structure, material structure, and social structure.

Distributed cognition returns culture, context, and history, to the picture of cognition. But these things cannot be added on to the existing model of cognitive processes without modifying the old model. That is, the new view of culturally embedded cognition requires that we remake our model of the individual mind in ways that incorporate social, cultural, and environmental structures as well as the wider ecologies of activity systems and processes involved in meaningful human activity.

Similarly, data about real-world activity becomes essential to evaluating models of cognition and ethnography, the fine-grained examination of real-world behavior, and plays an increasingly crucial role. The methods of participant observation and analysis of audio and video recordings are the stock in trade of the ethnographer. These methods build up representation of activity systems in real-world contexts. They address questions such as "What are people really doing?," "What are they trying to do?," "What reasons do they give themselves and each other for doing what they do?"

9.4 Cognitive Ethnography

As a basis for work in our laboratory we are developing a method that we call *cognitive ethnography*. The goal of cognitive ethnography is an improved functional specification for the human cognitive system. Cognitive psychology has traditions of testing hypotheses within particular research paradigms, but the relevance of these hypotheses to the activities that people actually engage in is largely unknown (Neisser 1982). Concerns about the ecological validity of experiments have fostered attempts to assess the relationship between experimental tasks and real-world tasks (Cole 1996). Yet, there is no way to address this relationship in the absence of a careful study of real-world cognition.

Historically, cognitive ethnography has fallen between the borders of the traditional disciplines of psychology, anthropology, sociology, communication, and linguistics. While each of these fields has an interest in either cognition or ethno graphy, until very recently ethnographic studies of cognition have been rare. The most significant initiative in this domain came in the 1960s and 70s when cognitive anthropology emerged as a distinct subfield. But there the theoretical framework of the studies was a sort of information processing psychology that saw cognition as a disembodied process (D'Andrade 1995). Because of these theoretical roots, ethnographic studies of cognition in this tradition ignored everyday activity (Werner and Schoepfle 1987).

A primary goal of cognitive ethnography is to better understand everyday activity.

Fine-grained examination of real-world behavior provides evidence counter to two traditional divides. First, an examination of cognition situated in interactions with the social and material world narrows the gap between thought and action. Doing is a kind of thinking and thinking is doing. Perception and action turn out to be more closely related than had previously been thought (Clark 2001; Noë 2003). Second, real-world activity always has an affective component. Cognition and emotion are more closely linked than has traditionally been assumed. The examination of real-world activity can reveal the nature of this linkage (Goodwin and Goodwin 2001). These considerations create a demand for ethnographic studies that focus on cognitive processes as they are enacted in naturally situated activity. We call this cognitive ethnography.

The theoretical emphasis from distributed cognition is reflected in a methodological focus on events. Since the dynamic cognitive properties of systems that are larger than an individual play out in the activity of the people in them, a cognitive ethnography must be an event-centered ethnography. We are interested not only in what people know, but in how they go about using what they know, to do what they do.

It is important to note that cognitive ethnography is not any single data collection or analysis technique. Rather it brings together many specific techniques, some of which have been developed and refined in other disciplines (e.g., interviewing, surveys, participant observation, video and audio recording). Which specific technique is applied depends on the nature of the setting and the questions being investigated. Because of the prominence of events and activity in distributed cognition theory, we give special attention to video and audio recording of events.

As alluded to earlier, a new generation of inexpensive digital recording devices and storage facilities is revolutionizing data collection in behavioral science, extending it into situations that have not typically been accessible and enabling examination of the fine details of action captured in meaningful settings. The ability to record and share such data has not only created a critical moment in the practice and scope of behavioral research but also presents unprecedented opportunities and challenges for the design of future augmented environments. In the next section, we briefly discuss examples of research projects attempting to meet these challenges and exploit opportunities presented by new technological facilities.

9.5 Examples of Research Projects: Lessons for Designing Augmented Environments

Here we sample five recent research projects from our laboratory. Each was chosen to exemplify a distributed cognition perspective and highlight opportunities and challenges relevant to the design of augmented environments. The first project demonstrates the richness of activity that can be captured by augmenting an automobile with ten cameras and discusses the opportunity for computer-vision techniques to assist the analysis of the video data. The second describes projects capturing histories of workstation activity and the analysis challenge the finding of patterns in this rich data. The third highlights the challenges and opportunities of using data from high-fidelity flight simulators to help to understand the complex socio-technical system of commercial aviation. The fourth describes the challenge of designing a multimodal augmented environment to assist communication between deaf patients and non-signing physicians. The fifth describes opportunities associated with creating paper and audio augmented digital documents.

9.5.1 Video Capture of Driving Activity: The Opportunity of Computer Vision to Reduce the Cost of Analysis

Computer vision techniques have advanced in capabilities and reliability to the point that they promise to be highly useful tools for aiding analysis of video data. To characterize this potential we describe recent experience automatically annotating video recordings of driving activities. The goal of this project is to understand the cognitive ecology of driving as the basis for designing instrumentation and controls to augment and improve driver safety (McCall et al. 2004; Boer et al. 2005). In order to ground design in real driving behavior, our lab in collaboration with Mohan Trivedi's Computer Vision and Robotics Research Lab instrumented an automobile to record multiple video streams and time-stamped readings of instruments and controls. This included video from a head-band 3rd-Eye camera we developed and from an array of other cameras (see Fig. 9.1) positioned to capture views from within and around the car, as well as of the driver's face and feet. Timeline-based representations were particularly useful for assisting with analysis of video of this rich activity data and for associating data from multiple instruments. Figure 9.2 depicts graphs of selected car parameters coordinated by time and linked to GPS-derived freeway locations.

Fig. 9.1 *Top*: Head-band "3rd Eye" camera (*left and middle*) and view from it (*right*). *Bottom*: Example views from selected other cameras (composite on *left*, omniview in *center*, and rear view on *left*)

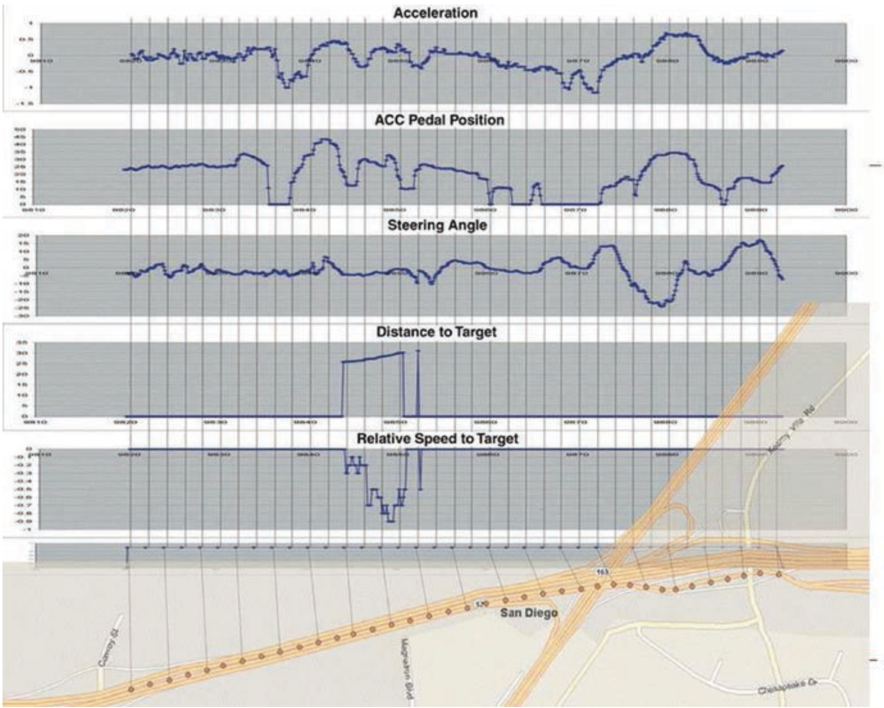

Fig. 9.2 Results from an analysis tool we developed to allow analysts to graph selected car parameters. This can include results from automated video analyses and can be linked by time or to GPS-derived freeway locations

We are encouraged by our success in automatically annotating this video data to aid analysis. For example, we developed code to compute the lateral angular velocity of the head from the 3rd-Eye camera video. This allows identification of even small head position adjustments as well as glances to the rear view mirror, glances to the left or right side mirror, and large over-the-shoulder head movements. By thresholding the amplitude of recorded audio we indexed times when someone was speaking in the car. Foot motion and lateral foot position were extracted from a "Foot-Cam" video using a simple detection algorithm. In combination with recordings of brake pedal pressure this easily enables determining, for example, when drivers move their foot to the brake in preparation for braking. We also developed a code to determine where the hands were positioned on the steering wheel and to automatically compute lateral position of the car as a basis for detecting lane changes. It is important to note that unlike most work in computer vision, annotating video to aid analysis does not typically require real-time processing. Offline processing is usually sufficient.

This work has led to an on-going effort in our lab to build an Ethnographers Workbench that integrates a variety of computer-vision and other facilities to assist annotation and analysis of video data. While we do not have space to review all the vision-based techniques we see as promising for automatic annotation, we briefly mention two examples: object recognition and face and emotion detection.

9.5.1.1 Object Recognition

It would be a boon to digital video analysis if the computer could automatically label all (or even most) frames or segments of a video in which a particular object is present. For example, if analysts are interested in activities involving interaction with specific objects, they might want to view only those segments of video that involve those objects. One very promising candidate technique uses distinctive invariant features extracted from training images as a basis for matching different images of an object. An important aspect of the Scale Invariant Feature Transform (SIFT) technique (Lowe 2004) is that it generates large numbers of features that densely cover the image over the full range of scales and locations. The features are invariant to image scale and rotation, and provide robust matching across a substantial range of affine distortions, additions of noise, and changes in viewpoint and illumination.

Since this algorithm is probabilistic, we can allow the user to modify the algorithm's threshold depending upon the task. For example, the threshold for object detection could be set at a low value in which virtually every frame that contains the object is detected, with the price of having some false alarms (flagged frames in which the object is not actually present). In this case, a small amount of user intervention would be required in order to cull the false alarms from the true detections. On the other hand, the object detection threshold could be set at a high value so there would be virtually no false alarms (every flagged frame is a true detection), with the price that in some frames the object would be present but not detected.

Depending upon the analysis task (finding every instance vs. finding a collection of representative instances), one or the other threshold (or somewhere in between) might be appropriate.

9.5.1.2 Face and Emotion Detection

There are myriad ways in which computerized face detection and face tracking could enable new types of analyses with huge potential gain and minimal time commitment on the part of the analyst. Current face detection algorithms (Viola and Jones 2002) could be employed to annotate the video so that appropriate video segments could be located quickly and accurately. One example of the state of the art in current research is work by Tim Marks, one of our recent Ph.D. students, and colleagues to automatically annotate a video with the subjects' emotions, as determined by their facial expressions (Marks et al. 2004). Computerized analysis of facial expression can be done with existing technology on the frontal views of faces. To analyze facial expressions from non-frontal views of a person's face, sophisticated 3D tracking algorithms such as G-flow (Marks et al. 2004) can be used to find the 3D pose of the face and the 3D locations of key points on the face from 2D video of the subject. By fitting the 3D locations of these key points to a database of laser scans of human heads (Blanz and Vetter 1999), we can synthetically rotate the face from any viewpoint to a frontal view, from which the emotion of the subject can be determined using a facial expression analysis system.

9.5.2 Activity Trails: Challenge of Finding Patterns in Workstation Activity Histories

There is a long history and a recent resurgence of interest in recording personal activity. Personal storage of all one's media throughout a lifetime has been desired and discussed since at least 1945, when Vannevar Bush published *As We May Think* (Bush 1945), positing the Memex, a device in which an individual stores all their books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. His vision was astonishingly broad for the time, including full-text search, annotations, hyperlinks, virtually unlimited storage and even stereo cameras mounted on eyeglasses.

Today, storage, sensor, and computing technology have progressed to the point of making a Memex-like device feasible and even affordable. Indeed, we can now look beyond Memex at new augmented environment possibilities. In particular, while media capture has typically been sparse throughout a lifetime, we can now consider continuous archival and retrieval of all media relating to personal experiences. For example, the MyLifeBits project (Gemmell et al. 2006) at Microsoft Research is recording a lifetime store of information about the life of Gordon Bell. This includes not only video but the capture of a lifetime's worth of articles, letters, photos, and presentations as well as phone calls, emails, and other activities. This and related projects are documented in the recent series of ACM CARPE workshops on capture, archiving, and retrieval of personal experiences.

Hollan has long been interested in visualizing activity histories. In early work on ReadWear and EditWear (Hill et al. 1992), he and his colleagues modified an editor to collect detailed histories of people editing text or code and made those histories available in the scrollbar of the editor in ways to inform subsequent activity. Over the last few years lab members have conducted a series of pilot projects collecting workstation activity of users. In one effort, Etienne Pelaprat built facilities to record a low-level operating system call activity on a workstation and then explored parsing that low-level activity record into higher level activity descriptions. The motivation was to have recording facilities that did not require modification to any applications participants used in their normal workstation activity.

Gaston Cangiano is currently building a system, Activity Trails (Cangiano and Hollan 2009), to unobtrusively capture histories of all user activity on the desktop. The tool has very low memory overhead and employs low level system calls so that there is no performance impact noticeable to users. The tool records a bitmap snapshot of the desktop at regular intervals, as well as XML output containing all information about input devices, windows and applications (including titles and content for selected applications). Studies in cognitive ethnography often suffer from uneven sampling so that it is impossible to know how representative any particular event is of overall behavior patterns. By collecting a complete record, Activity Trails supports the creation of frequency distributions for the various categories of observed events.

ActivityTrails also has the capability to playback summaries for given times or dates. A screenshot of the design is shown in Fig. 9.3. The playback area at the bottom allows users to scroll back and forward in time. The thumbnails on top represent individual episodes. The time elapsed between each episode varies depending on the results from parsing the recordings. One goal is to provide an episodic view in which thumbnails index landmarks or natural breakpoints in user activity.

There is considerable evidence that people organize their memory for activities around landmarks (Robinson 1986; Tulving 2002; Shum 1994; Huttenlocher and Prohaska 1997). The challenge is to develop a parsing algorithm to detect breakpoints in activity that match what people report as the boundaries of their activities. We have been capturing workstation activity from workers in a law office and from undergraduate students. To obtain subjects' views of workstation activities we have had them explain what they were doing as they watch screen recordings of their past activities. We are testing how well parsing algorithms we have been developing can match the activity boundaries that participants mention. While results are still preliminary, we are encouraged by the ability for cross-correlations of mouse, window, and keyboard activity to determine boundaries. The fact that people develop stable windowing styles for different types of activities is a nice example of how cognitive activity is distributed through time. Stable windowing practices facilitate the allocation of attention in subsequent activity, thus reducing cognitive workload.

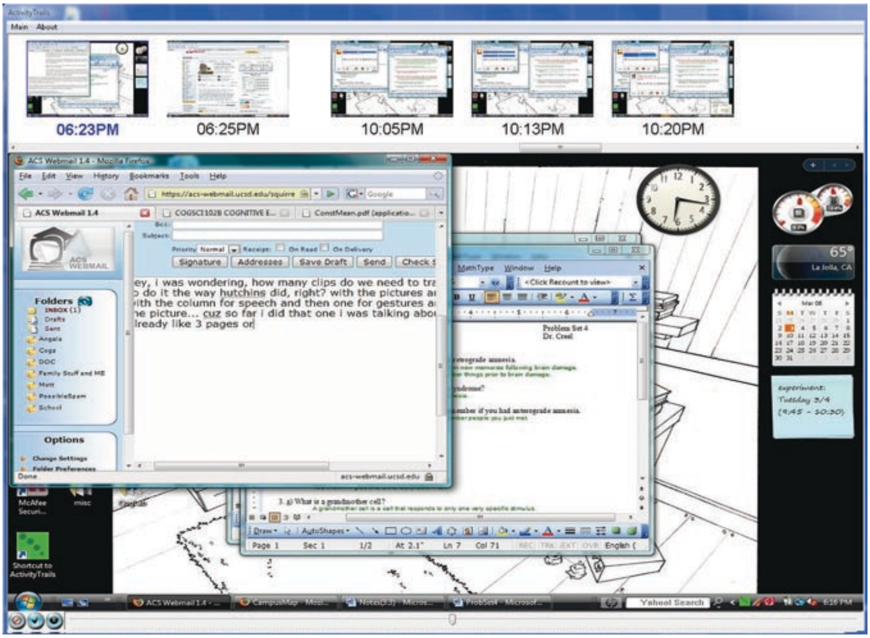

Fig. 9.3 Episodic view of past activity – summaries: each thumbnail in the *top* portion of the window represents a short activity

9.5.3 Commercial Aviation: Challenge and Opportunities of Understanding a Complex Socio-Technical System

Under a multi-year agreement with the Boeing Commercial Airplane Group, Hutchins has negotiated access to training activities at a number of airlines outside the US (Security provisions enacted in the wake of the 9/11 terrorist attacks have made work with US airlines impossible). Boeing's interests lie in specific applications and interventions concerning training, operating procedures, and flight deck design in the next generation of airline flight decks. Hutchins has already collected video data in Japan, New Zealand, Australia, and Mexico. In addition to collecting data in simulators located at the training centers of non-US airlines, the project also collects data in simulators located at the Alteon/Boeing training center in Seattle.

These data provide an absolutely unique look at complex, highly structured, expert activity in a setting that is spatially, temporally and institutionally constrained. The activities recorded in high-fidelity flight simulators, a most interesting augmented environment, are complex. They involve the production of multimodal acts of meaning making that are embedded in social and material context. The script-like structure of the phases of flight and of flight deck procedures provides a common framework with respect to which of the activities of different pilots and even different populations of pilots can be compared. This attribute of the activity also makes it a good early choice for exploration with timeline representations because there are recognizable shifts in activity structure in successive phases of flight. By the time they are in training for commercial air operations, pilots have high levels of expertise. Studying expert real-world skills is important, but difficult to do because analysts must have considerable expertise themselves in order to interpret the significance of the presence (or absence) of particular behaviors. Fortunately, Hutchins' years of experience as a jet-rated pilot and as an aviation researcher provide the necessary analytic expertise. The spatial and temporal constraints on activity in the flight deck make data collection tractable in the sense that recording equipment can be installed in fixed locations, or attached to the clothing of the participants, and the activities to be recorded are sure to take place in an anticipated amount of time (usually about 2 h). The institutional constraints guarantee that the data recordings will be rich in observable activity (little down time) because of the high cost of operating the simulator. Finally, as was the case with our instrumented car, it is possible to acquire a rich digital data stream from the simulator itself. Time synching simulator data to the observational data provides a documentary richness that is simply not possible in most activities.

Our observations in Japan have already revealed that language practices in the Japanese airline flight deck can be seen as adaptations to a complex mix of exogenous constraints. Institutions, such as regulatory agencies, adapt to the constraints of global operations when setting the rules that govern airline operations; the decision that air traffic control communications shall be conducted in English, for example. Airlines adapt to the regulatory environment, the marketplace, the characteristics of their workforce and the nature of the technology when setting training and operational policies. Pilots adapt to the residues of the adaptive behaviors of institutions in constructing meaningful courses of action in flight (Hutchins et al. 2006).

Commercial aviation is a complex socio-technical system that has developed most rapidly in North America and Europe. Because we are working with non-US airlines, we are also able to examine how other cultures integrate the practices of commercial aviation into their particular cultural and cognitive ecology (Nomura et al. 2006). This is a unique perspective on the globalization of one of the most complex socio-technical systems in today's world. Dramatic changes in the demographics of the global population of commercial airline pilots are currently underway. For example, the mean age of pilots worldwide is rapidly decreasing as aviation expands in Asia and Latin America.

The data collected in the Alteon/Boeing training center is especially interesting because it involves American flight instructors working with pilots from other nations. These data permit us to examine the dynamics of a special case of intercultural learning. Airline pilots everywhere share certain elements of professional culture, but in intercultural training, professional culture becomes a resource for overcoming the boundaries of national culture. They permit us to see the contextual grounding of intercultural communication and learning (Hutchins et al. 2006).

9.5.4 A Multimodal Augmented Environment: The Challenge of Supporting Medical Conversations Between Deaf and Hearing Individuals

Loss of hearing is a common problem that can result from noise, aging, disease, and heredity. Approximately 28 million Americans have significant hearing loss, and of that group, almost six million are profoundly deaf. A primary form of communication within the United States deaf community is American Sign Language. It is estimated to be the fourth most commonly used US language. While ASL is widely used in the US, no one form of sign language is universal. It is important to note that ASL is not just a visual form of English; it is a different language with its own unique grammatical and syntactical structure.

ASL interpreters play a central role in enabling face-to-face communication between many deaf and hearing individuals. For the deaf population fluent in ASL, communicating through an interpreter is an optimal choice for many scenarios. Interpreters, however, are expensive and not always available. Furthermore, though interpreters are bound by a confidentiality agreement, the presence of a third person in a highly private conversation may reduce a deaf person's comfort and inhibit their willingness to speak candidly.

While other viable communication tools for the deaf community exist, in our lab Anne Marie Piper and Hollan have been exploring tabletop displays with speech recognition to facilitate medical conversations between deaf and hearing individuals (Piper and Hollan 2008). Consultations with physicians often involve discussion of visuals such as medical records, charts, and scan images. Interactive tabletop displays are an effective tool for presenting visual information to multiple people at once without necessarily designating one person as the owner of the visual. Taking notes while meeting with a physician is problematic for deaf individuals because it requires simultaneously attending to the doctor's facial expressions, the interpreter's visual representation of speech, and notes on paper. A tabletop display allows all the active participants to maintain face-to-face contact while viewing a representation of the conversation in a central location. The Shared Sound Interface Piper and Hollan designed incorporates keyboard input by the patient and speech input by the doctor, allowing the physician to speak and gesture as they discuss medical details and visuals with the patient. SSI leverages the affordances of multimodal tabletop displays to create an augmented environment to enhance communication between a doctor and a patient, potentially transforming a challenging situation into a constructive and collaborative experience (Fig. 9.4).

SSI uses a MERL DiamondTouch table (Dietz and Leigh 2001) and the DiamondSpin toolkit (Shen et al. 2004). The DiamondTouch table is a multi-user, multi-touch top projected tabletop display. Users sit on conductive pads that enable the system to uniquely identify where each user is touching the surface. SSI enables conversational input through standard keyboard entry and a headset microphone. The audio captured from the microphone is fed into a speech recognition engine (currently this is Microsoft Windows' default recognizer, but the application easily

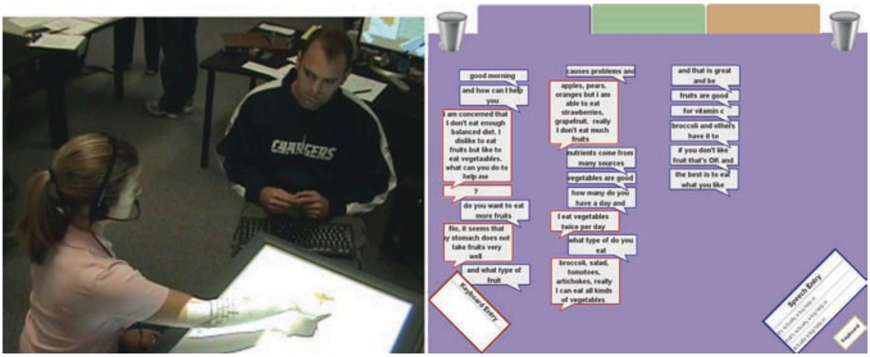

Fig. 9.4 Doctor (*left person*) and patient (*right person*) point together as they discuss part of a medical visual (*left*). A screen shot of the shared speech interface application (*right*) shows that speech bubbles from a doctor and patient consultation persist on the display and are available for manipulation and later reference

adapts to any off-the-shelf recognizer). SSI uses the Java Speech API and CloudGarden software to interface with the speech recognition engine and send converted speech-to-text into the main application running on the DiamondTouch table.

To evaluate the SSI system Piper and Hollan conducted a laboratory study with eight deaf participants who were born deaf or became deaf before the age of one. All participants were fluent in ASL and proficient at reading and writing in English. Each deaf participant conversed with a medical doctor and a professionally trained ASL interpreter about a sample medical issue. Overall, participants indicated that digital tables are a promising medium for facilitating medical conversations. Survey data indicated that the application was good for private conversations and enabled independence.

There were several key differences in communication between the Digital Table condition and a condition in which an ASL interpreter was present. The Digital Table condition allowed for asynchrony in communication, whereas the interpreter acted as a broker of conversation and thus encouraged synchronous interactions. Conversation in the Interpreter condition was the fastest and allowed a greater number of turn taking exchanges. Equitable participation levels were observed in the two conditions, the doctor and patient each contributed to about half of the conversation.

In the Digital Table condition, non-verbal and gestural communication played an important role in augmenting speech and ensuring successful communication. The co-located, face-to-face nature of the digital table allowed participants to provide feedback to their partner about their state of understanding through deictic gesture (e.g., pointing), gaze sharing, and head nodding. There were also differences in how patients attended to the doctor when the interpreter was present. Deaf participants looked at the doctor when they signed but then shifted their gaze to the interpreter when the doctor began speaking. In the Digital Table condition, participants typically looked at the doctor when she was speaking and then looked down at the display.

From the perspective of augmented environments, it is important to note that the digital table transformed the ephemeral nature of speech into a tangible and persistent form, thereby creating affordances that are not available in traditional conversation. For example, there were interesting behaviors with the speech bubbles because of their form. When a phrase was added to the display that referred to a previous utterance, the "owner" of the speech bubble often moved the new phrase close to the previous utterance. In conversation, the speaker must help listeners understand a reference to a previous utterance through context and explicit referencing. The digital table allowed users to reference previous conversation by placing new speech near an existing speech bubble. Similarly, the doctor and patient used the tail of the speech bubble as a pointing mechanism. That is, participants strategically placed speech bubbles around the display so that the tail of the speech bubble pointed to part of a background visual. The persistent nature of speech with the digital table allowed participants to review their conversation. Both the doctor and patients looked back over their previous conversation. The doctor said "it was good to look back at what I had covered with that particular patient," and explained that "it would be helpful because it is not uncommon in medicine to have very similar conversations with different patients throughout the day."

This work exemplifies the importance of adopting a unit of analysis that spans across the doctor and patient as well as their speech, gestures, and interactions with the digital table. Understanding the functional relationships among these elements is fundamental to designing an effective augmented environment.

9.5.5 Paper and Audio Augmented Digital Documents: Opportunities from Bridging the Paper-Digital Divide

Paper persists as an integral component of virtually all environments and tasks because it provides ease of use unmatched by digital alternatives. Paper documents are light to carry, easy to annotate, rapid to navigate, flexible to manipulate, and robust to use in varied environments. Interactions with paper documents create rich webs of annotation, cross reference, and spatial organization. Unfortunately, the resulting webs are confined to the physical world of paper and, as they accumulate, become increasingly difficult to store, search, and access. XLibris (Schilit et al. 1998) and similar systems address these difficulties by simulating paper with tablet PCs. While this approach is promising, it suffers not only from limitations of current tablet computers (e.g., limited screen space) but also from loss of invaluable paper affordances.

For the last several years, Hollan has been collaborating with Francois Guimbretière and others on the design of PapierCraft a gesture-based command system that allows users to manipulate digital documents using paper printouts as proxies (Liao et al. 2008). Using an Anoto digital pen (Anoto 2008), users can draw command gestures on paper to tag a paragraph, email a selected area, copy selections to a notepad, or create links to related documents. Upon pen synchronization, PapierCraft executes the commands and presents the results in a digital document viewer. Users can then search the tagged information and navigate the web of annotated digital documents resulting from interactions with the paper proxies. PapierCraft also supports real time interactions across mix-media, for example, letting users copy information from paper to a Tablet PC screen.

Recently we have started to explore a new Anoto-based technology that supports the integration of audio. The Livescribe Pulse pen (Livescribe 2008) functions like earlier digital pens but also provides the ability to capture and play audio. For example, during a meeting one can record what is being said while taking notes on paper. Subsequently, touching a section of one's paper notes allows one to hear what was being said at the time that section of notes was being written. This technology enables a variety of interesting opportunities to augment environments.

One opportunity we are pursuing is to incorporate Livescribe pens into the Ethnographer's Workbench we are designing. For an ethnographer, the ability to easily capture audio and index into it via associated paper notes changes the relations between listening and writing in powerful ways. The content of written notes can shift from recording speech to coding and other sorts of meta-data representations of the recorded event. We also are exploring the ability to create special data notebooks that provide the ability to flexibly annotate sections of notes by touching categories printed on the pages. Recently, we initiated the design of such a notebook for a research team studying synchrony of behavior between a Beluga whale and her new calf. The ability to automatically record the temporal aspects of the data is particularly valuable. Similarly, timestamps can be used to link paper and audio notes with camera images or video. The ButterflyNet system (Yeh and Klemmer 2005), designed to support field biologists is an interesting example.

9.5.6 Lessons for Designing Augmented Environments

A rapidly evolving technology landscape and recent advances in cognitive theory make this an exciting time for the design of augmented environments. We see great potential for exploiting new facilities to record the fine details of action captured in meaningful settings as a basis for better understanding of the dynamics of human behavior and thus providing a solid foundation for designing augmented workplaces in which people can pursue their goals in collaboration with the elements of the social and material world.

In this chapter we have briefly described what we see as the major challenge for behavioral science: developing the requisite theoretical and methodological frameworks to effectively analyze data that capture the richness of real-world human activity. We then discussed selected recent projects in our laboratory in which we are attempting to further develop distributed cognition and cognitive ethnography. We conclude by characterizing lessons our work has for designing and evaluating augmented environments:

1. There is currently a tremendous opportunity to exploit computer vision techniques to automatically annotate video data and thus assist analysis and evaluation. Today the high labor cost of analyzing rich activity data leads to haphazard and incomplete analyses or, all too commonly, to no analysis at all of much of the data. Even dataset navigation is cumbersome. Data records are chosen for analysis because of recording quality, interesting phenomena, and interaction density – producing a haphazard sampling of the recorded set. Good researchers have a nose for good data, but also have a tendency to focus on small segments of the record that contain "interesting" behavior, analyze them intensively, and then move on to the next project. When analysis is so costly, few analyses can be done – so datasets are severely underutilized – and researchers come to have a large investment in the chosen data segments. Since each analysis may appear as an isolated case study, it can be difficult to know how common the observed phenomena may be. Larger patterns and contradictory cases can easily go unnoticed. Well-known human confirmation biases can affect the quality of the science when each analysis requires so much effort. Thus, one promising research strategy is to develop and assemble tools and practices to increase speed and improve analysis. Here we argue that computer vision techniques can be employed to automatically annotate video data and help to manage and coordinate data analysis.
2. Since so much of peoples' activity currently takes place while using computers, there is an unprecedented opportunity to capture and study this activity. This enables not only examination of detailed activity histories to better understand the dynamics of cognition but also can serve as the basis for summaries of those activities to help people to reestablish previous contexts and find needed information. Augmenting environments to encapsulate histories of activity and provide visualizations of it is an important research direction. We see it being useful for augmented environments like we described (workstations, driving, flight simulators, and multimodal interfaces) as well as for the increasing range of ubiquitous computing environments currently being developed.
3. *Tools to help capture, visualize, and analyze activity data are fundamental to advancing research*. We argue that integrated capture, analysis, and visualization facilities such as we are designing can increase speed, improve, and help coordinate analyses focused at different scales so as to profit from the complementary perspectives of multiple disciplines. For example, the Ethnographer's Workbench we are building incorporates paper and audio augmented digital notebooks with facilities to support multiscale visualization and analysis of video and other time-based data. A significant scientific challenge for all disciplines is how to represent data so as to make important patterns visible. In the behavioral sciences, researchers transcribe and code data in a wide variety of ways, creating new re-representations of the original events. Currently the coordination of multiple re-representations with the original data is typically done by hand, or not at all. Since this re-representation process – including all sorts of transcription, coding system development and implementation, and re-description – is what allows us to do scientific work, even small improvements in automating coding, transcription, or coordination of representations can be crucially important. Recent developments

> in behavioral science theory create special challenges in this regard. Increasingly theories are concerned with patterns that can emerge from the interactions of many dynamically linked elements. Such interactive patterns may be invisible to approaches that decompose behavior into the more or less independent components created by historical distinctions among behavioral science disciplines. This is why multidisciplinary behavioral science is necessary. But tools that match this multidisciplinary vision are also needed.

The richly multimodal nature of real-world human activity makes analysis difficult. A common strategy has been to focus on a single aspect of behavior or a single modality of behavior, and to look for patterns there. However, the causal factors that explain the patterns seen in any one modality may lie in the patterns of other modalities. In fact, recent work suggests that activity unfolds in a complex system of mutual causality. Analysis may still be based on decomposition of the activity, as long as there is a way to put the pieces back together again. That is, as long as there is a way to visualize the relations among the many components of multimodal activity.

The structure of the existing academic disciplines attests to the fact that human behavior can be productively described at many levels of integration. Neuroscientists describe regularities at a finer scale than psychologists, who describe phenomena at a finer scale than linguists, who in turn tend to describe behavior at a finer scale than anthropologists. A deep understanding of the nature of human behavior demands not only description on multiple levels, but integration among the descriptions.

As behavior unfolds in time, describable patterns that take place on the scale of milliseconds are located in the context of other describable patterns that display regularities on the scale of seconds. Those patterns in turn are typically embedded in culturally meaningful activities whose structure is described on the scale of minutes or hours. Patterns at larger time scales are created by and form the context for patterns at shorter time scales. Visualizing and reasoning about such nested temporal relations requires representations that allow coordination of analyses across multiple scales.

References

Anoto (2008) http://www.anoto.com

Bauman R, Sherzer J (1989) Explorations in the ethnography of speaking. Cambridge University Press, Cambridge

Blanz V, Vetter T (1999) A morphable model for the synthesis of 3D faces. In: SIGGRAPH 99 Proceedings, pp 187–194

Boer E, Forster D, Joyce C, Fastrez P, Haue JB, Garvey E, Choski M, Mogilner T, Hollan J (2005) Bridging ethnography and engineering through the graphical language of petri nets. In: Measuring Behavior; 5th International Conference on Methods and Techniques in Behavioral Science Research, Wageningen, The Netherlands

Brooks R (1991) Intelligence without representation. Artificial Intelligence Journal 47: 139–159

Brown JS, Collins A, Duguid P (1989) Situated cognition and the culture of learning. Educational Researcher 18(1): 32–42

Bush V (1945) As we may think. The Atlantic Monthly 101–108

Cangiano G, Hollan J (2009) Capturing and restoring the context of everyday work: a case study of a law office. In: Proceedings of HCI International. In press.
Card S, Mackinlay J, Shneiderman B (1999) Readings in information visualization: using vision to think. Morgan Kaufmann Publishers, California
Chaiklin S, Lave J (1996) Understanding practice: perspectives on activity and context. Cambridge University Press, Cambridge
Clancy WJ (1997) Situated cognition: on human knowledge and computer representations. Cambridge University Press, Cambridge
Clark A (1997) Being there: putting brain, body and world together again. MIT, Cambridge
Clark A (2001) Mindware: an introduction to the philosophy of cognitive science. Oxford University Press, New York
Clark A (2003) Natural-born cyborgs: minds, technologies and the future of human intelligence. Oxford University Press, New York
Cole M (1996) Cultural psychology: a once and future discipline. Belknap Press, Massachusetts
D'Andrade RG (1995) The development of cognitive anthropology. Cambridge University Press, New York
Dietz P, Leigh D (2001) DiamondTouch: a multi-user touch technology. In: Proceedings of the 14th Annual Symposium on User Interface Software and Technology. ACM, New York, pp 219–226
Dourish P (2001) Where the action is: the foundations of embodied interaction. MIT, Cambridge
Gemmell J, Bell G, Lueder R (2006) MyLifeBits: a personal database for everything. In: Communications of the ACM. ACM, New York, pp 88–95
Goldin-Meadow S (2003) Hearing gesture: how our hands help us think. Belknap Press, Cambridge
Goodwin C (2000) Gesture, aphasia, and interaction. In: McNeill D (ed) Language and gesture: window into thought and action. Cambridge University Press, Cambridge
Goodwin C, Goodwin M (2001) Emotion within situated activity. In: Duranti A (ed) Linguistic anthropology: a reader. Blackwell, Massachusetts, pp 239–257
Goodwin C, Heritage J (1990) Conversation analysis. Annual Review of Anthropology 19: 283–307
Gumperz JJ, Hymes D (1986) Directions in sociolinguistics. Blackwell, Oxford
Hill W, Hollan J, Wroblewski D, McCandless T (1992) Edit wear and read wear. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, p 3–9
Hollan JD, Bederson BB, Helfman J (1997) Information visualization. In: Landauer TK, Prabhu P, Helander MG (eds) The handbook of human computer interaction (Chap. 2). Elsevier Press, Amsterdam, pp 33–48)
Hollan J, Hutchins E, Kirsh D (2000) Distributed cognition: Toward a new theoretical foundation for human-computer interaction research. In: ACM Transactions on Human-Computer Interaction. TOCHI '00. ACM, New York, pp 174–196
Hutchby I, Wooffitt R (1998) Conversation analysis: principles, practices and applications. Polity Press, Cambridge
Hutchins E (1995a) Cognition in the wild. MIT, Cambridge
Hutchins E (1995b) How a cockpit remembers its speed. Cognitive Science: A Multidisciplinary Journal 19: 265–288
Hutchins E, Klausen T (1996) Distributed cognition in an airline cockpit. In: Engestrom Y, Middleton D (eds) Cognition and Communication at work. Cambridge University Press, New York, pp 15–34
Hutchins E, Nomura S, Holder B (2006) The uses of paper in commercial airline flight operations. In: Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work. ACM, New York, pp 249–258
Huttenlocher J, Prohaska V (1997) Reconstructing the times of past events. In: Stein N, Orenstein P, Tversky B, Brainerd C (eds) Memory for everyday and emotional events. Lawrence Erlbaum, New Jersey
Jones V, Jones M (2004) Robust real-time face detection. International Journal of Computer Vision 57(2): 137–154

Lakoff G, Johnson M (1999) Philosophy in the flesh: the embodied and its challenge to western thought. Basic Books, New York
Lave J (1988) Cognition in practice. Cambridge University Press, Cambridge
Liao C, Guimbretière F, Hinckley K, Hollan J (2008) Papiercraft: a gesture-based command system for interactive paper. In: ACM Transactions on Computer-Human Interaction. ACM, New York, pp 1–27
Livescribe (2008) http:.//www.livescribe.com
Lowe D (2004) Distinctive image features from scale-invariant keypoints. In: International Journal Computer Vision. Springer Netherlands, pp 91–110
Marks T, Hershey J, Roddey C, Movellan J (2004) 3D tracking of morphable objects using conditionally gaussiean nonlinear filters. IEEE Conference Computer Vision and Pattern Recognition 12: 190
McCall JB, Achler O, Trivedi M, Fastrez P, Forster D, Haue J, Hollan J, Boer E (2004) A collaborative approach for human-centered driver assistance systems. Proceedings of the 7th IEEE Conference on Intelligent Transportation Systems 3: 663–667
McNeill D (2005) Gesture and thought. University of Chicago Press, Chicago
Nardi B (1996) Context and consciousness: activity theory and human-computer interaction. MIT, Cambridge
Neisser E (1982) Memory observed: remembering in natural contexts. WH Freeman, San Francisco
Noë A (2003) Action in perception. MIT, Cambridge, MA
Nomura S, Hutchins E, Holder B (2006) The uses of paper in commercial airline flight operations. In: Proceedings of Computer Supported Cooperative Work 2006 (CSCW 2006), pp 249–258
Nunez R (1996) Could the future taste purple? Reclaiming mind, body, and cognition. In: Nunez R, Freeman WJ (eds) Reclaiming cognition: the primacy of action, intention and emotion. Imprint Academic, Thorverton
Pea RD (1993) Practices of distributed intelligence and designs for education. In: Salomon G (ed) Distributed cognitions. Cambridge University Press, Cambridge
Piper A, Hollan J (2008) Supporting medical conversations between deaf and hearing individuals with tabletop displays. In: Proceeding of Computer Supported Cooperative Work 2008 (CSCW 2008), ACM, New York, pp 147–156
Port R, van Gelder T (1995) Mind as motion: explorations in the dynamics of cognition. MIT, Cambridge, MA
Prevignano CL, Thibault PJ (2003) Discussing conversation analysis: the work of Emanuel A. Schegloff. John Benjamins Publishing Company, Amsterdam
Robinson J (1986) Temporal reference systems and autobiographical memory. In: Rubing D (ed) Autobiographical memory. Cambridge University Press, Cambridge, pp 159–188
Rogoff B (2003) The cultural nature of human development. Oxford University Press, New York
Schilit B, Golovchinsky G, Price M (1998) Beyond paper: supporting active reading with free form digital ink annotations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, pp 249–256
Shen C, Vernier F, Forlines C, Ringel M (2004) DiamondSpin: an extensible toolkit for around-the-table interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, pp 167–174
Shum M (1994) The role of temporal landmarks in autobiographical memory processes. Psychology Bulletin 124: 423–442
Suchman LA (1987) Plans and situated actions: the problem of human-machine communication. Cambridge Press, Cambridge
Thelen E, Smith LB (1994) A dynamic systems approach to the development of cognition and action. Bradford Books, Cambridge
Thompson E (2007) Mind in life. Harvard University Press, Cambridge, MA
Tulving E (2002). Episodic memory: from mind to brain. In: Annual Review of Psychology. Annual Reviews
Varela FJ, Thomson E, Rosch E (1991) The embodied mind: cognitive science and human experience. MIT, London

Viola P, Jones M (2002) Fast and robust classification using asymmetric adaboost and a detector cascade. In: Advances in Neural Information Processing System. MIT Press,14:1311–1318

Werner O, Schoepfle GM (1987) Systematic fieldwork. Sage, California

Yeh R, Klemmer S (2005) ButterflyNet: mobile capture and access for biologists. In: Conference Supplement to UIST 2005: ACM Symposium on User Interface Software and Technology. ACM, New York

Zhao W, Chellappa R, Phillips PJ, Rosenfield A (2003) Face recognition: A literature survey. In: ACM computing surveys. ACM, New York, pp 399–458