SPECIAL ISSUE: Humans and Automation: Next Steps

# An Integrative Approach to Understanding Flight Crew Activity

**Edwin Hutchins**, **Nadir Weibel**, **Colleen Emmenegger**, and **Adam Fouse**, University of California, San Diego, and **Barbara Holder**, Boeing Company, Seattle, Washington

In this paper, we describe an integrative approach to understanding flight crew activity. Our approach combines contemporary innovations in cognitive science theory with a new suite of methods for measuring, analyzing, and visualizing the activities of commercial airline flight crews in interaction with the complex automated systems found on the modern flight deck. Our unit of analysis is the multiparty, multimodal activity system. We installed a variety of recording devices in high-fidelity flight simulators to produce rich, multistream time-series data sets. The complexity of such data sets and the need for manual coding of high-level events make large-scale analysis prohibitively expensive. We break through this analysis bottleneck by using our newly developed integrated software system called ChronoViz, which supports visualization and analysis of multiple sources of time-coded data, including multiple sources of high-definition video, simulation data, transcript data, paper notes, and eye gaze data. Four examples of flight crew activity serve to illustrate the methods, the theory, and the kinds of findings that are now possible in the study of flight crew interaction with flight deck automation.

**Keywords:** human automation interaction, topics, ethnography, methods, air transportation, ATC, domains

## INTRODUCTION

We focus on human–automation interaction located in the airline flight deck. This activity setting has structure, is dynamic, and emerges from complex interactions among many elements, including pilot knowledge, other pilot cognitive processes (such as attention, memory, embodied conceptualization, and expert motor skills), the flight deck interfaces, the behavior of automated systems, airspace structure and air traffic control (ATC), weather, and terrain. Typically, examining this activity involves moment-by-moment detailed transcription of pilots' speech, but a full understanding of the complexity of crew activity requires much more than an analysis of the crew's verbal behavior. The following transcript is from some data we collected at a training center belonging to a major air carrier in North America. The crew was flying a normal takeoff from San Francisco International Airport in the Boeing 737-300.

ATC: Oceanic 815 cleared for takeoff runway 2-8 right
PM: Oceanic 815 cleared for takeoff runway 2-8 right
PF: Brakes are released
PM: Thank you
PF: Going up together
PF: Takeoff [thrust]
PM: [OK] power set
PM: 80 knots power set
PM: V1 . . . VR . . . V2

The takeoff of a commercial airliner is a highly scripted multimodal ensemble performance that requires close coordination between the two pilots. In this flight, the captain (left seat) had the role of pilot monitoring (PM), and the first officer (right seat) was pilot flying (PF). (These roles imply a division of labor between the two pilots. Captain and first officer usually alternate taking the two roles on successive flight legs. The PF is responsible for controlling the airplane and ensuring that it goes where it is supposed to go. The PM is responsible for operating the aircraft systems, for communicating with air

Address correspondence to Edwin Hutchins, Department of Cognitive Science, University of California San Diego, 9500 Gillman Dr., La Jolla, CA 92093-0515, USA, ehutchins@ucsd.edu.

*Journal of Cognitive Engineering and Decision Making*
Volume 7, Number 4, December 2013, pp. 353–376
DOI: 10.1177/1555343413495547

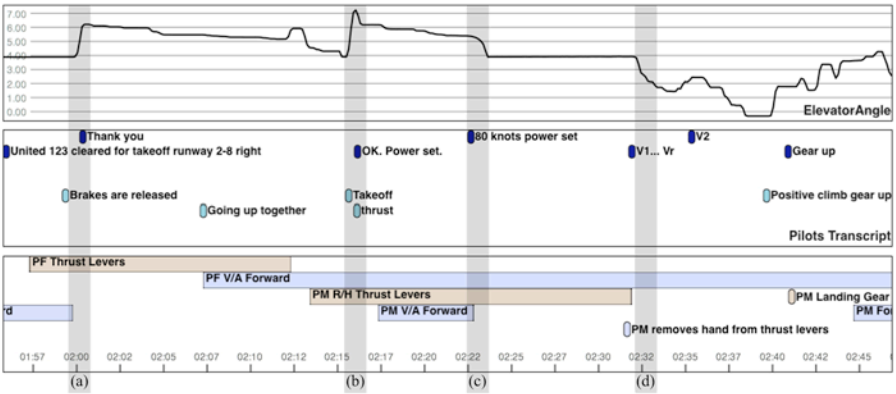



*Figure 1.* Three timelines showing the relationship between elevator angle plotted as a line graph in the top timeline (up on the graph is nose-down elevator) and the verbalizations of the pilot monitoring (PM; dark blue boxes) on the second timeline. On the bottom line are annotations about the tactile and visual attention of the two pilots. Neutral elevator is 4.0 units. Notice that the pilot flying (PF) applies nose-down elevator as soon as the brakes are released (a), then eases a bit until the PM declares, "OK, power is set" (b). PF holds nose-down elevator until PM says "80 knots power set," at which point he relaxes to neutral elevator (c). Shortly after PM says "V1 . . . VR," PF pulls the yoke back to lift the nose off the runway (d).

traffic control, and for monitoring both the PF and the situation of the aircraft.)

Our analysis will focus on three systems of coordination that are embedded in the takeoff activity: (a) the closed-loop control of the airplane with respect to the runway centerline achieved through the coordination of visual and proprioceptive/haptic attention of the PF, (b) the coordination of elevator input with airspeed as mediated by the visual and verbal attention of the PM and the auditory and proprioceptive/haptic attention of the PF, and (c) the distribution of labor between the PF and the PM in manipulating and "guarding" the thrust levers.

With the airplane in position on the takeoff runway, the PM handled communications with ATC listening to, and reading back, the takeoff clearance. Before the PM had finished reading back the takeoff clearance, the PF placed his left hand on the thrust levers and began pushing them forward to increase thrust. Both pilots visually monitored the engine instruments as the engines spooled up. To understand the rest of the story, we turn to three systems of coordination.

*Staying on the centerline:* The takeoff roll involves fine-scale coordination among several of the PF's attention modalities. The PF directs visual attention to the runway centerline and coordinates that visual attention with simultaneous proprioceptive and haptic attention to the rudder pedals, which steer the airplane. This task is a closed loop of negative feedback control to keep the airplane on the centerline. Visually monitoring the relationship of the aircraft to the runway centerline is demanding. The PF keeps his eyes mostly on the runway and relies on the PM to visually monitor the engine instruments and especially the indicated airspeed.

*Pitch control during the takeoff roll:* The division of labor with respect to visual attention is the foundation for the control of pitch inputs by the PF. While advancing the thrust levers, the PF pushed the control yoke forward. He then relaxed the forward pressure a little while the thrust was coming up. When the PM said, "Thrust is set," the PF pushed the yoke forward again. Just after the PM said, "80 knots," the PF relaxed all forward pressure on the yoke. When the PM said, "V1 . . . VR," the PF pulled back on the yoke and lifted the nose wheel off the runway. The PF must pull back on the yoke at rotation speed to increase the angle of attack and generate the lift required to begin flying. This coordination for pitch control is illustrated as three timelines of activity in Figure 1.

A reader may wonder why the PF makes pitch inputs when the airplane is on the ground. To understand the relation between pitch inputs and callouts by the PM during the takeoff roll, a little background is required. Once thrust is at takeoff level, the failure of an engine on either wing would put a huge yawing force on the airplane. If this yaw is not countered immediately by a steering command, the airplane could go off the side of the runway. Below about 80 knots, there is insufficient aerodynamic force in the rudder to counter this asymmetric yaw. Nose-wheel steering would have to counter the yaw, but in order to be effective, the nose wheel must be firmly in contact with the runway. Thus, the airplane is especially vulnerable to a runway excursion in the period from the time thrust reaches takeoff power on both engines until the time that speed reaches about 80 knots. Nose-down elevator (produced by pushing forward on the yoke) increases the downward pressure on the nose wheel during the takeoff roll. That is why the PF pushes the yoke forward during the takeoff roll in coordination with the verbal cues provided by the PM that bracket the period in which preserving the integrity of nose-wheel steering is most important.

*Handling the thrust levers:* A third system of coordination involves the thrust levers. Thrust levers control engine power; pushing the levers forward increases thrust, and pulling them back decreases thrust. The PF advanced the thrust levers for the two engines together and visually monitored the engine instruments to ensure that the two engines increased in thrust symmetrically. In this section of the transcript, he commented, "Going up together." Once he verified that both engines were spooling up normally, he engaged the autothrust system by pushing a button on the thrust levers. The autothrust system automatically moves the thrust levers forward to a position that will produce a precomputed amount of thrust for the takeoff. The PF allowed his left hand to ride the thrust levers forward until they were nearly at the takeoff setting. He then removed his hand from the thrust levers, and the PM smoothly took control of the thrust levers, positioning his right hand on them with his fingers dangling down in front of the levers. The PM's hand remained in that position until he said "V1," at which time he tapped the levers once and moved his hand away.

In this sequence of actions, the PF initiated the setting of takeoff thrust, and the PM (the captain) kept his hand on the thrust levers until V1. V1 is the "decision speed" prior to which the captain has the authority to reject the takeoff if something goes wrong. To do so, he would call out, "Abort," and pull the thrust levers to idle. After V1, the airplane is committed to takeoff because not enough runway remains to bring the airplane to a safe stop. Even if an engine fails after V1, the airplane will continue with the takeoff. To help the crew to avoid rejecting the takeoff with insufficient runway remaining, the captain removes his hand from the thrust levers when the V1 speed is reached.

This description should make it clear that in order to really understand what the crew is doing, we would have to somehow capture the activity in many modalities, including speech, gesture, visual and haptic attention, manipulation of controls, and airplane state. We would also have to re-represent the raw sensor data as context-dependent phenomena relevant to the flying activity and somehow visualize the relationships among the phenomena. Capturing the behavior of this complex system requires a network of sensors and recording devices. The move to richer recordings of ongoing behavior is under way and will continue. Sensor technology is changing rapidly and a very wide variety of inexpensive sensors is now available. In addition to capturing the output of sensors that directly sense some aspect of the crew behavior, researchers typically take notes while observing. Doing so allows the researcher to note high-level processes, such as coordination among members of a crew. The means to measure all of these aspects of crew activity are currently available. That is the good news. Measuring and analyzing activity in all of these modalities implies huge and rich time-based data sets. The bad news is that coding and re-representing raw sensor data as events of interest in a human factors analysis is prohibitively expensive. These costs create a bottleneck in the analysis process. Most detailed analyses of flight crew activity are "one-off" endeavors. These analyses are useful for illustrating how such systems work, but they do not support generalization and comparison.

In this paper, we report our efforts to break through the analysis bottleneck and sketch out a future of human–automation studies. To avoid being overwhelmed by such volumes of data, we have developed a suite of computational tools that accelerate the analysis of data and enable the assessment of complex phenomena. We make use of multiple high-definition video cameras to capture activity at different granularity levels and from different perspectives. We equip each pilot with a microphone and high-fidelity digital audio recorder. We track the eye gaze of both pilots in a flight deck using portable eye tracking glasses. We collect log files and sensor data from fixed-based or motion-based flight simulators. We use digital pen and paper both to record the pilots' interactions with paper material and to augment the observers' (e.g., researchers' or instructors') note-taking practices. Finally, we create a unified view of the collected data for interacting with and analyzing this information. We believe that this novel suite of methods for computation and analysis opens up exciting avenues for the definition of new approaches to research and training.

## Developments in Cognitive Science Theory

At the core of our approach is the theory called distributed cognition and a suite of methods called cognitive ethnography. In the past two decades, we have seen dramatic changes in cognitive science theory. The field is moving from a concept of cognition as a logical mechanism to the concept of cognition as a biological process. This shift has been driven in part by the rapid expansion of cognitive neuroscience but also by the increasing realization that thinking happens in the interaction of the body with an environment for action and not just in the brain. Cognition is increasingly viewed as a process that extends beyond the skin and skull of the individual (Clark, 2003; Cole, 1996; Hollan, Hutchins, & Kirsh, 2000; Hutchins, 1995a; Nardi, 1996; Norman, 1993; Pea, 1993; Suchman, 1987). This shift in framing the unit of analysis for cognition introduces a host of previously overlooked cognitive phenomena to be documented, studied, and understood. The fields of embodied cognition (Gibbs, 2006), extended mind (Clark, 2008), and distributed cognition (Hutchins, 1995a) have demonstrated the value of a new unit of analysis for cognitive phenomena that includes the bodies and the world of the cognitive agents.

Because these approaches expand the boundaries of the system in which cognitive processes occur, they also indicate new places to look for cognitive phenomena as well as new kinds of cognitive phenomena. This shift has a place in our efforts to understand interactions between human operators and automated systems. As our unit of analysis, we take the *flight deck system*, composed of two pilots in interaction with each other, the airplane, and the airspace via the interfaces provided in the flight deck. Our objective is to develop a new understanding of human–automation interaction in terms of processes that unfold in the complex flight deck system. Our foci are on this system rather than on the individual pilot, on the full range of cognitive processes rather than on mental states alone, and on multimodal perception and action loops rather than on visual perception alone.

## The Flight Simulator as a Training and Data Collection Environment

High-fidelity flight simulators are powerful tools for pilot flight training, providing pilots the opportunity to practice flight operations and to encounter and respond to dangerous situations, all in a safe environment. We believe that flight simulators could be even more valuable training tools in the future than they are today. Many flight simulators are equipped with simulator log file and video recording capabilities, but video recordings of training sessions are often not used in the debriefing following a simulator session. We believe that much can be done to provide tools that better exploit existing capabilities and introduce new approaches to training.

Simulators are also a great boon to aviation human factors researchers. A high-fidelity flight simulator is a well-developed virtual reality environment. Simulators provide researchers the ability to observe and record crew behavior in conditions that are close to the real world while still permitting some control of events through the design of scenarios. A simulator can be instrumented with a variety of data collection devices, including video cameras, audio recorders, eye trackers, and so on. It is important to notice also that the simulator itself can serve as an important data collection apparatus. Through the traces they leave in the simulator log files, the control yoke, rudder pedals, thrust levers, trim switches, and other controls become devices that allow us to measure aspects of proprioceptive and haptic attention.

However, as with research in any rich and meaningful setting, simulator-based research on flight crew activity always encounters the analysis bottlenecks described earlier. For both flight instructors and researchers, the data that are so easily collected in a simulator are not easily accessed. A video recording of an hour or two in length is not a welcoming object of analysis. Searching a video recording (even a digital video) for a particular important event is a time-consuming process. For example, if an instructor wanted to review a particular maneuver with student pilots, the instructor would first have to find the event in question in the video recording. If a researcher wanted to examine all of the instances in which a crew member made a mode selection on the Mode Control Panel (MCP), the researcher would have to search the video carefully, noting the temporal locations of the events of interest. Modern video transcription applications, such as Noldus Observer (Zimmerman, Bolhuis, Willemsen, Meyer, & Noldus, 2009), Mangold Interact (http://www.mangold-international.com/), ELAN (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006), VCode (Hagedorn, Hailpern, & Karahalios, 2008), Transana (http://www.transana.org), and Diver (Pea et al., 2004), provide tools to facilitate annotation and coding of video recordings, but the search of the video is still a manual task for the operator.

Imagine a world in which the instructor or researcher could, immediately upon the conclusion of a simulator session, navigate the video recordings via an annotated timeline that labeled and displayed the temporal locations of all of the events of interest. Using a new data visualization tool called ChronoViz, we have constructed such a world.

## ACTIVITY IN CONTEXT

In this section, we describe the integration of the measures of activity across the modalities and also show how the meanings of the actions of the crew as well as the behavior of the automated systems are established in the context of the flight operations activities. None of the modalities of activity (visual, auditory/verbal, proprioceptive/haptic) can be fully constructed from a single data stream. Even eye gaze data must be supplemented by other (video or audio) data to be accurately reconstructed. The eye tracker can give us ($x$, $y$) coordinates of gaze, but to know what the pilot was attending to, when looking at the given location, requires additional data.

Furthermore, the meaning of actions, regardless of modality, can be understood only in the context of their occurrence because meanings are properties of larger units of activity; the meanings of the parts are derived from the whole rather than the other way around (Hutchins & Nomura, 2011; McNeill, 2005). Similarly, the meaning of the actions of a single pilot is not recoverable from the actions of that pilot alone, nor are the meanings even in the actions and interactions of the two pilots. Rather, the meanings are in the organization of the activity of the entire flight operations system, including the pilots, automated systems, airplane, airspace, and ATC.

### Progressions of Complexity

We have chosen a set of brief analyses of flight deck activities to illustrate the use of our methods. Each example will be described in several ways:

- As an activity in the domain of flight deck operations (takeoff, setting the heading bug, display management, flight crew–automation interaction)
- As an example of one or more phenomena in the domain of cognitive processes (distribution of attention, visual to verbal transduction, crew communication, consequential action)
- As a set of time-series data streams produced by recording devices (over-the-shoulder video, eye-tracker scene camera video, audio, simulator data, eye tracking coordinates)
- As a set of analysis procedures applied to the data streams in ways that allow us to visualize the phenomena of interest (event recognition in simulator data, dual eye tracking analysis, recurrence plot, eye gaze to object registration via infrared markers, and via computer vision)

**TABLE 1:** Activity Table

| Activity | Phenomena | Time Series | Analysis |
|---|---|---|---|
| Coordinating Activity | | | |
| Takeoff: Manual flight | Multiparty, scripted multimodal interactions; distribution of attention, crew communication | Multiple videos, audio, simulator data | Manual analysis with ChronoViz; transcription, automatic event recognition and performance measures |
| Heading select: Simple autoflight usage | Microdynamics of multimodal attention | Video, eye tracking | Manual analysis with ChronoViz; examination of dynamics of eye and hand coordination |
| Multimodal Meaning Making | | | |
| Navigation: Display management | Multiparty micro-multimodal meaning making; speech, gesture, action | Video, audio, eye tracking | Manual analysis with ChronoViz; audio transcription, gesture transcription |
| Route modification: Crew–automation interaction | Multiparty, multimodal interaction with aircraft automation | Video, audio, eye-tracking | Manual analysis with ChronoViz; coordinated eye gaze, gaze recurrence |

These examples are shown in Table 1. They are arranged in order to produce a progression of increasing complexity with respect to each description. The activities progress in complexity of flight operations from manual flight to the selection of a precompiled arrival procedure to be executed by the autoflight system. The phenomena progress from simple coordination of macroscopic action to complex acts of multiparty multimodal meaning making. The time-series data progress from video, audio, and simulator data to the inclusion of eye gaze and then dual eye gaze. The analysis progresses from manual coding of observable actions to automatic event recognition in the simulator data stream to the analysis of the coordination of eye gaze and gaze recurrence. The remainder of the paper will illustrate and explain these examples through an analysis of the related phenomena and modalities.

## Analyzing Multimodal, Multiparty Activity Data

To study activity as an integrated system, we need tools that allow us to look at multiple sources of data about the activity in a synchronized and integrated way. We have developed such a tool as part of our research. This tool, called ChronoViz (Fouse, Weibel, Hutchins, & Hollan, 2011), supports visualization and analysis of multiple sources of time-coded data, including multiple sources of high-definition video, simulation data, transcript data, paper notes, and eye gaze data. Each data source can be independently aligned with the rest of the data and then used for navigation of the data set as a whole. Researchers using ChronoViz have rich interactive capabilities for exploring data sets, making annotations about observed activity, filtering and arranging annotations, and performing computational analysis of the loaded data.

Figure 2 shows the basic ChronoViz interface as used to visualize data collected during a flight simulation. Synchronized video files at the top play together with the use of standard video controls. Below the videos are a number of timelines, showing annotations about the data and line graphs of simulator variables. The movement of an index bar across the timelines is

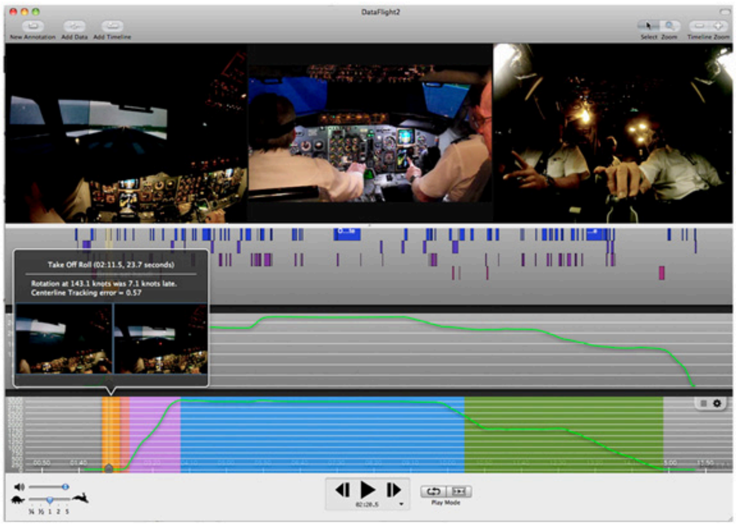



*Figure 2.* A basic ChronoViz session with three synchronized videos in the top row, a visualization of the speech transcript in the top timeline (different colors represent utterances by different speakers), the airspeed of the aircraft in knots in the center timeline, and the aircraft pressure altitude graphed as a line in the bottom timeline. The colored regions on the bottom timeline are the automatically computed phases of flight. The popup window displays data about the takeoff roll phase of flight, including the duration of the phase, the airspeed at rotation, the difference between actual and planned speed at rotation, and a measure of accuracy of centerline tracking.

synchronized to the video files. Any representation of the data can be used to navigate the entire data set. One can find a video frame of interest and then read the displayed simulator values or examine the transcribed utterances. Similarly, one can click on the timeline to move the index bar to the location where a graphed simulator variable reaches a particular value and view the activity in the videos for that moment during the simulation.

ChronoViz can also read the latitude and longitude values from the imported simulator data and display the path of the airplane in geographic space (see Figure 3a). Annotations that have been made about the data, such as the phases of flight shown in Figure 2, can be overlaid on the path to help the researcher navigate the data set. This representation is also synchronized with the other data streams. The dot on the flight path represents the location of the airplane at the selected moment. Moving to a new point in any of the other representations of the data automatically repositions the dot. Clicking on the dot, and dragging it along the flight path, will cause all of the other data representations to update to the time at which the airplane was at the selected position in the flight path.

The entire data set can also be navigated in ChronoViz via interaction with the transcript of the audio recording. At present, our transcripts, shown in Figure 3b, are produced by trained human transcribers. The transcription is done in the InqScribe application (http://www.inqscribe.com), then imported into ChronoViz. In the future, when technology becomes accurate enough, speech-to-text algorithms tailored to the aviation domain and its technical terms could also be used instead of transcribers.

While ChronoViz can display heterogeneous time-based data sets in a synchronized way, it can also generate derived data combining any of the variables that have been imported into a session. The automatic identification of flight phases and the performance measures within each phase, shown in Figures 2 and 3a, illustrate this key general feature of ChronoViz. Any analysis that can be computed on the data streams imported into ChronoViz can be used to generate annotations on the timeline representations. In this case, we created analysis programs (plug-ins) that analyze the numeric data stream from the simulator in order to locate, label, and annotate the phases of flight. The space of possible

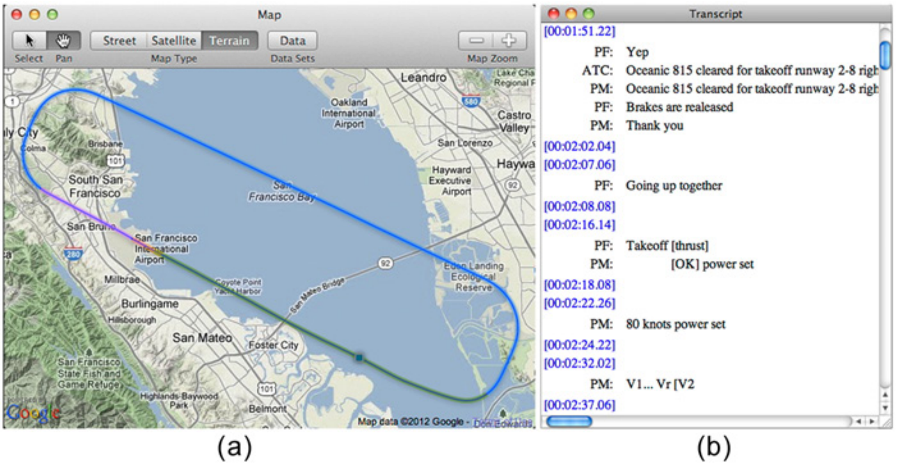



*Figure 3.* Geographic information and transcribed speech shown in ChronoViz.

analyses is large and can be customized to users' needs. Analysis is not limited to the simulator data stream. Other analysis programs could be applied to the verbal transcript or even to computer vision analysis of the contents of video streams. An analysis program might draw data from more than one data stream. For example, an analysis program to automatically determine the response latency between a wind shear alert and the first move to execute the escape maneuver might draw on simulator data as well as on computer vision analysis of a video stream.

ChronoViz also supports integration of notes taken with a digital pen (Weibel et al., 2012; Weibel, Fouse, Hutchins, & Hollan, 2011) by exploiting the Anoto Digital Pen and Paper technology (http://www.anoto.com). This capability is used to augment the observational abilities of researchers or instructors as well as to include notes taken by pilots as part of the integrated data collection. The digital pen data is integrated in the same way as the other types of data. Researchers can click on notes to move to the point in the data when those notes were taken, and ChronoViz visualizes the progression of the notes by highlighting corresponding notes as the video is played back. Notes can also be interactively explored by selecting regions of notes to reveal a timeline representation of when those notes were recorded. This capability is especially useful when taking notes that are organized spatially, as in the annotated diagram shown in Figure 4.

## COORDINATION OF ACTIVITY

Our first two examples emphasize the nature of the coordination of activity. When we say that a collection of elements is a system, we are referring to the dynamic patterns of coordination among the elements. This definition is true whether the system is a single human being or a group of people interacting with one another and an activity setting.

We opened the paper with a description of a normal takeoff in the Boeing 737-300. While developing the analysis of the takeoff example, we utilized the basic capabilities of our system: multiple video streams, digital audio, and numeric data from the simulator. The coordination and analysis of these data sources produced the timelines shown in Figure 1. One of the video streams was from a head-mounted camera. Head orientation is a reasonable proxy for eye gaze when the areas of interest (AOIs) are large and separated in space. This is the case for judgments about when the PM was looking at the engine instruments and when he was looking through the windscreen at the runway.

The fact that a takeoff is a closely orchestrated ensemble performance is not a discovery. However, in the context of research or training, being able to measure, quantify, visualize, and

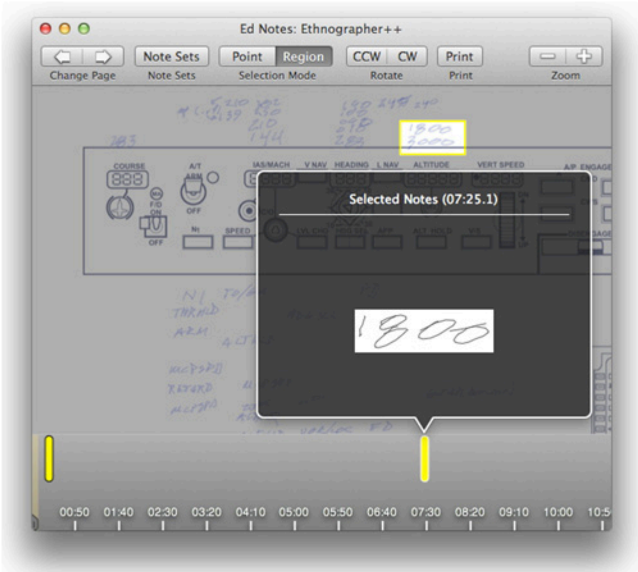



*Figure 4.* A ChronoViz window showing digital notes in the form of an annotated diagram. Notes were recorded on a printed Mode Control Panel diagram with a digital pen during observation of a flight simulator session, then imported into ChronoViz. Annotations made about the altitude control are selected, and the timeline below the notes shows when those annotations were made. Hovering the mouse over the marks on the timeline shows detail about the note.

analyze the fine details of the flight crew activity can be very valuable. It can allow us to answer questions such as the following: How is the coordination among crew members and between crew and aircraft actually achieved? What attentional resources are recruited by the activity? How are those resources organized? And how does such a system adapt to perturbations? In this example, we presented a normal takeoff. Imagine a similar analysis of a takeoff that includes a problem, such as an engine failure, or a more subtle problem, such as an unreliable airspeed reading. Mapping the allocation of flight crew attention in such a scenario could inform the design of alerts. We might even take the takeoff as a model for pilot monitoring in other operational contexts.

Our analysis of the takeoff focused on the dynamic coordination between the airplane interface and the behaviors of the members of a flight crew. Researchers, instructors, and pilots often create explicit representations of this sort of interpersonal coordination of macroscopic action patterns. The coordinated joint action patterns of the takeoff are script-like. They are represented in training manuals and can be taught to pilots.

In the second example (presented next), we examine a simple automation management action: changing the position of the heading bug while in heading select mode. Of course, pilots can be trained to use the interface to change the position of the heading bug. However, here we focus on the coordination among aspects of different modalities of attention within a single crew member. Such patterns of intrapersonal coordination among microscopic action patterns are not available to consciousness and have no conventional representation in language or in training materials. This sort of coordination is opportunistic and is not trainable by conventional means. However, once it is understood, it is likely that these patterns of allocation of multimodal attention can be sculpted through the incorporation of appropriate design features in the interface to the automated systems. Furthermore, we suspect that some planned design interventions may inadvertently interfere with virtuous aspects of the current flight deck interface system.

*Figure 5.* Boeing 787 Dreamliner flight simulator.

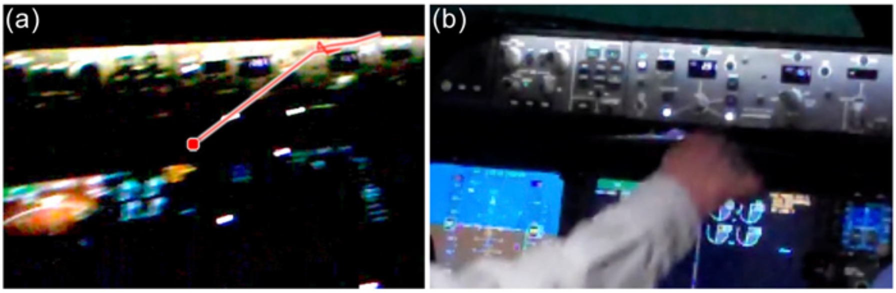


*Figure 6.* Synchronized frames from the eye tracker scene camera and the over-the-shoulder video showing hand–eye coordination during heading selection. The overlay line in panel (a) shows the history of eye gaze over the previous two seconds.

## Heading Select

The takeoff procedure described earlier took place in a simulated Boeing 737-300. The Boeing 737-300 is an old airplane and is much less automated than contemporary models. In order to get a better understanding of flight deck activity in a state-of-the-art commercial airliner, we collected data on two qualified pilots flying a simulator for Boeing's new 787 Dreamliner (B787). In this simulation event, the PF was in the left seat and the PM was in the right seat.

One of the most basic autopilot functions is the tracking of a heading target. The autoflight mode that accomplishes this is called *heading select*. The crew sets the heading target to be tracked by turning the Heading Select (HDG SEL) knob located on the MCP, which is an autopilot interface panel built into the flight deck glare shield. Figure 5 shows the PF and PM in the B787 simulator during our data collection session (note that the pilots are wearing eye tracking glasses). The figure also shows a detailed view of the MCP. Using the HDG SEL knob to specify a heading target is a frequently occurring task. It involves reaching for and grasping the control knob, then rotating it to select the desired heading. The selected heading is displayed numerically as three digits in a window above the knob and also graphically as a heading index (also known as the heading bug) attached to the end of a dashed line on the navigation displays (NDs).

Many instances of setting a heading occurred during our B787 simulator session. Let us first describe an entirely routine instance of a heading change that illustrates the close relationship of the visual, proprioceptive/haptic, and tactile modalities in fine eye–hand coordination.

In order to analyze the relationship of the involved modalities, and to reason about the resulting coordination, we performed a multimodal microanalysis of the event. We made use of two cameras (one over-the-shoulder high-definition camera and a scene camera of the PF's eye tracking device) and the recorded eye gaze position and carefully studied the event on a micro scale, concentrating our analysis on three modalities and their coordination: (a) tactile, (b) proprioceptive, and (c) visual. Figure 6 presents synchronized frames from the PF's eye tracking

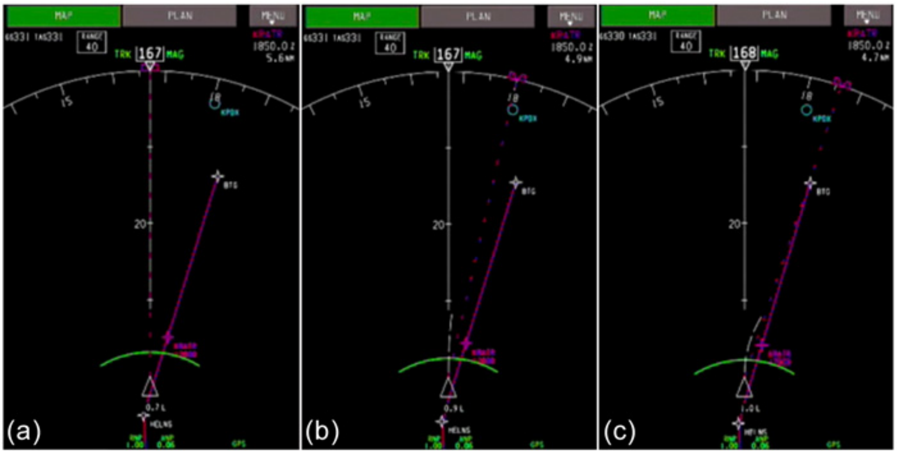



*Figure 7.* The pilot flying's navigation display before the heading bug was moved (a), while he was moving the heading bug (b), and after he was done moving the heading bug (c). The current position of the heading bug is indicated by the dashed magenta line.

glasses scene camera on the left and the over-the-shoulder video on the right.

During much of the flight, the PF's eye gaze was on the primary flight display (PFD), while his right hand was resting on the thrust levers, where he could monitor changes in commanded thrust through proprioception without looking at the engine instruments. While in that position, the PF noticed a condition on the PFD that required a change in heading: The airplane had departed from its intended route.

Upon noticing this departure, his right hand and eyes set out toward the HDG SEL knob simultaneously (Figure 5a). As might be expected, his eye gaze arrived at the HDG SEL knob, establishing the location of the grasp target before his hand arrived. Interestingly, as shown in Figure 6, which shows a moment approximately 1 s after his gaze first reached the HDG SEL knob, the PF's eyes departed back to the PFD while his right hand was still approximately 6 in. from the HDG SEL knob. Visual attention was required to establish the location of the grasp target, but it was not necessary to keep looking at the knob while it was grasped because of the physical nature of the knob (Figure 5b). The tactile experience of the HDG SEL knob is unlike anything in its vicinity. Fine visual input is not required to guide the grasp once the hand is on course to make contact with the knob (proprioceptive mode). However, guidance and feedback are needed in order to know how to manipulate the knob to select the desired target heading. Although the PF could have looked at the window just above the knob displaying the heading as a numerical value, his eye gaze went instead to the ND, where the PF followed the movement of the heading bug, as illustrated in Figure 7. The PF turned the HDG SEL knob with his right hand while monitoring the effects of that manipulation by visually tracking the heading bug on the ND.

The nature of the task and the configuration of the displays explain the coordination. The important information to the PF is the location of the dashed magenta line shown on Figure 7, with respect to the selected waypoint (the white BTG [Battle Ground] point on the right of the display). Although it is difficult to tell what numerical value of the heading would take the airplane to the waypoint it is quite easy to see the dashed magenta line move into position over the waypoint while the HDG SEL knob is being manipulated. The relevant information to the pilot is therefore not a number in the HDG SEL window but a spatial relationship on the ND. This function looks simple but is actually an example of the powerful properties offered by a graphical interface to the autoflight system.

Interestingly, this same coordination pattern was not found in other tasks that required heading changes using the HDG SEL knob. When the nature of the task changes, the coordination pattern changes. For example, we observed how during the approach to the destination airport, the ATC cleared the airplane to fly a specific set of headings (e.g., 90° turn from downwind—280°—to base—190°). In this activity, the numerical description of the new heading is the relevant information. In this case, the PF's eye gaze and right hand set out to meet at the HDG SEL knob, but the eye gaze remained in the vicinity of the HDG SEL knob while the new heading target was selected.

Multimodal coordination is involved in virtually all flight deck tasks, but the way the modalities are coordinated depends heavily on the specifics of the activity. The same task (selecting a heading) but with different triggers (aligning with a direction given by visual angle vs. relative change of the heading in degrees) is accomplished with a different configuration of attention modalities (hand, eye, proprioceptive). Although those subtle differences could be observed only through a multimodal microanalysis, we feel that they are of paramount importance to fully understand the dynamics of multimodal attention.

These observations have implications for contemporary trends in flight deck design. Many dedicated hardware devices are now being replaced by touch-sensitive flat-panel displays. The last example demonstrated how in certain situations, fingers can negotiate the details of the grasp without visual input. If the current MCP were replaced by a touch-sensitive flat-panel display, additional visual resources would be required to get the finger on the correct location. This requirement would slow down the process and introduce additional possibilities for error. In fact, we wonder if these results have already happened in the replacement of the old control display unit (CDU), located between the pilots just ahead of the thrust levers. The old CDU has been replaced with a touch-sensitive virtual CDU. Is it possible that pilots could exploit the tactile structure of the line-select keys on the old mechanical CDU in ways that freed visual attention for other useful purposes? If so, then savings somewhere else in the system (e.g., reduced maintenance costs for flat-panel displays) is passed along to the pilots in the form of an unmeasured human performance cost.

As flight deck technology changes, designers face difficult trade-offs. On the one hand, virtual displays and touch screens allow for display flexibility and may have lower maintenance costs than fixed hardware switches and knobs. On the other hand, our examination of the microscopic details of allocation of attention shows that interactions with shape-coded hardware knobs can exploit features of the human multimodal attention system in ways that make action more effective and robust. We can now measure previously unmeasured (perhaps previously ignored) aspects of this important design trade-off.

This heading select example illustrates the interaction between a pilot and a basic autopilot function. In this example, we expanded our data collection devices to include a wearable eye tracking system. When fine distinctions of angle of regard are required, head orientation is not sufficient to determine direction of gaze. The additional richness of the data allows us to see an unexpected relationship in the system of hand–eye coordination. This relationship is possible only in the environment of dedicated shape-coded hardware controls. Were the MCP to be replaced by a touch-sensitive flat-panel display (as has been done with the CDU), then some efficiencies in the current ecology of attention allocation could no longer be achieved.

## MULTIMODAL MEANING MAKING

In the previous section, we presented an example of coordination among crew members and an automated system as well as an example of coordination of multiple modalities of attention between a single crew member and an automated system. In every case of interpersonal coordination, it is true that intrapersonal coordination is also involved. In this section, we address cases that involve the *measurement and analysis* of both interpersonal and intrapersonal coordination.

### Display Reconfiguration

As noted earlier, one of the advantages of virtual displays over dedicated hardware units is flexibility in display configuration. All of the principal displays and some controls in the B787 can be moved around the instrument

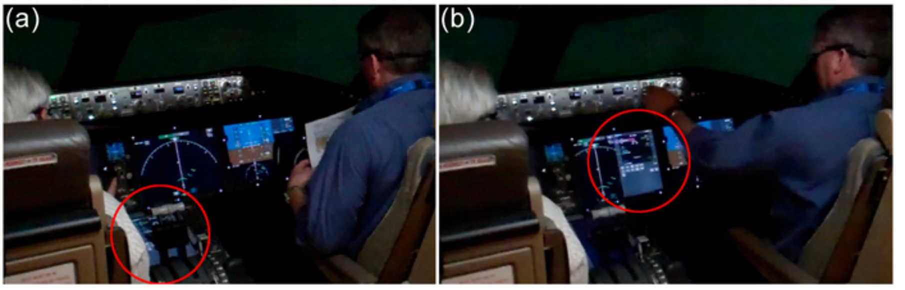



*Figure 8.* Boeing 787 Dreamliner flight simulator instrument panel showing the control display unit and navigation display positions.

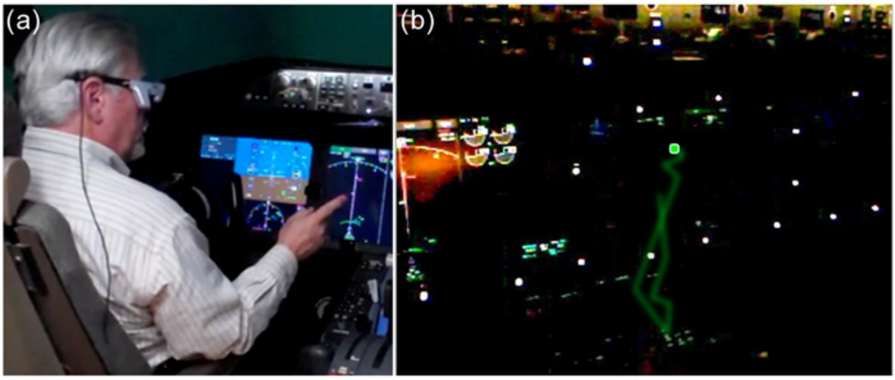
panel in response to failures and also to enhance flight crew workflow. Display reconfiguration is a relatively new activity on the flight deck. The CDU is the primary interface to the flight management system. Through this interface, which incorporates a small screen and a keypad, the crew enters and views information about the performance of the airplane, navigation parameters, and the flight route. As shown in Figure 8a, the default location of the virtual CDU in the B787 is on the front edge of the center console, in the location where the dedicated hardware CDU was located on earlier Boeing airplanes. Because the virtual CDU operates via touch-sensitive regions rather than hardware switches, it can be moved to a variety of locations on the instrument panel. For example, the CDU can be relocated to the main instrument panel on the right side of what is normally the first officer's ND, as shown in Figure 8b.

Electronic checklists have recently replaced paper checklists in modern airplanes. Electronic checklists have an advantage over paper in that the checklist system can interrogate the state of the airplane and can set the status of some checklist items automatically. The electronic checklists can also be displayed in place of the CDU. Reconfiguration of the displays is accomplished by manipulating a display management panel on the glare shield. The actions required to reconfigure the display are interesting, but they are not the focus of our discussion here. Rather, we want to answer questions such as "Why would a flight crew decide to reposition the CDU?" and "How does a flight crew accomplish the repositioning?" In order to understand the "why" and "how" of CDU repositioning, we performed a multiparty micro-multimodal analysis of the event, looking at the coordination of visual, tactile/proprioceptive, and auditory/speech modalities within the crew and between the crew and the aircraft's automated flight system.

Before taxiing the airplane in the simulated Boeing 787 flight, the PM called up the preflight checklist. This checklist was displayed in the default CDU position on the right side of the center console, just to the left of the PM's knee. In this location, the checklist did not respond to the PM's attempts to check off completed items on the list. The PM said, "OK, for some reason that checklist is not working on that screen." He continued with, "So let me try up here." The phrase "up here" refers to a position on the main instrument panel, directly in front of the PM, in the right half of the PM's ND (see Figure 8b). After moving the checklist to the right side of his ND, the PM found that he could successfully interact with the checklist both using a cursor and by touch.

We can determine the referent of the indexical phrase "up here" because in our integrated multimodal data set, it can be seen that just before this utterance, with his left hand still on the virtual buttons of the lower checklist, the PM's eye gaze made a quick saccade up to the ND (the referent of "up here"), briefly back down to the checklist, and then up to the display reconfiguration controls. His hand also began moving up when he began this utterance fragment. Because of the multiparty nature of our data, we also know that the PF watched this event unfold.

About 7.5 min after the successful use of the repositioned preflight checklist, the crew encountered a problem with the CDU. The flight




(c) [00:35:36.16]

PM: [See it's not letting me manipulate] that either.

PF: [Still not taking it?]

[00:35:39.12]

PF: Will it work on your... your screen up there, like the uh check list?

PM: Uhhhh you know maybe. Let's see ... so if we put CDU up there...

[00:35:50.21]

*Figure 9.* Multimodal coordination of gesture, eye gaze, and talk by the pilot flying (PF) during the attempted arrival menu selection. His communication with the pilot monitoring (PM) was shaped by their previous interactions with the preflight checklist and their shared physical space. The PF's eye gaze was initially on the preflight checklist displayed on the center console. The PF's eye gaze then moved briefly up to the PM's navigation display (ND) on the forward instrument display, where he later proposed the control display unit be reconfigured. He then moved his eye gaze back down to the current location of the checklist. The images in panels (a) and (b) were captured as the PF looked to the PM's ND.

was in its cruise phase, and the crew was working to complete the construction of the flight route. In particular, the crew was attempting to select an arrival procedure at the destination that would bridge the gap between the en route airways and the instrument approach to the destination runway. Using the CDU in its default position on the center console, the PM was unable to display the menu of arrival procedures for the destination airport. The PM used his eye gaze to locate touch targets on the virtual CDU. The PF's eye gaze was also on the CDU when the PM depressed the ARR button to bring up the arrivals menu. When the display did not change, the PM said, "See it's not letting me manipulate that either." At the same time, the PF asked, "Still not taking it?" (Figure 9c). The PF's eye gaze then shifted up to the right side of the PM's ND, on the main instrument panel (Figure 9b). At this time, the panel was displaying only the ND. This location is where they had earlier successfully interacted with the preflight checklist. About half a second after moving his eye gaze to the PM's ND, the PF gestured, pointing up to the location he was looking at. After another half second of gesturing and looking, the PF asked, "Will it work on your . . . your screen up there, like the uh check list?" The PM replied, "Uhhh you know maybe. Let's see . . . so if we put CDU up there," and then quickly reconfigured the display location.

Using the eye tracking data, the audio data, and gesture analysis from the video data, we know that the PF was closely monitoring the

PM's activity at this time. The analysis shows he was aware of the PM's verbal expression of frustration, as did his body's expression of frustration (pressing the menu select button multiple times) at not being able to activate the menu. The PF's body was turned slightly right, toward the location of the activity, and his eye gaze was on the PM's ND. Given his training and experience in advanced commercial airplanes, we know that the PF knew the CDU could be repositioned to the ND. Because of the checklist activity described earlier, we know that the PF also knew that the checklist, which did not work in the default position, did work when positioned on the main instrument panel. We know that the intended referent of his pointing gesture with his right hand (see Figure 9) was the PM's ND, because the gesture was part of a three-part multimodal utterance. The utterance included the spoken words "Will it work on your . . . your screen up there, like the uh check list?" as well as the eye gaze directed to the right side of the PM's ND and the pointing gesture.

Although none of these elements alone is unambiguous, together they *mutually elaborate* one another (Goodwin, 1994). The meaning of the indexical gesture is established by its relation to the PF's eye gaze and spoken words. The referent of the indexical words "up there" is established by its relation to the gesture and the eye gaze. And finally, the meaning of the eye gaze for the speaker, PF, is enacted in its relation to his gesture and his words. It is worth noting here that eye gaze is not simple intake of information; it is an action on the world. This assertion will become even more obvious in the next example.

This analysis demonstrates that reciprocal monitoring by the flight crew is accomplished through multimodal behavior. Not only does the PM monitor the activities of the PF, but the PF also monitors the activities of the PM. The ways that the PF monitors the PM are not well studied in aviation human factors. This analysis also demonstrates joint reasoning, with the crew making inferences about the similarity of the CDU to the electronic checklist and an inference about how the functioning of an interface might be dependent on the position in which it was displayed. Lacking any explanation as to why the CDU would not allow the pilots to display the menu of arrival procedures, and having just witnessed the preflight checklist work in the upper instrument panel and not the lower center console, they decided to see if the CDU would work properly in the upper display location. The analysis also makes clear the value of collecting data over periods long enough to cover a set of activities and not just one activity in isolation. As we saw, events that occurred some minutes earlier in the activity provide context for the interpretation by the researcher of the current activity of the crew.

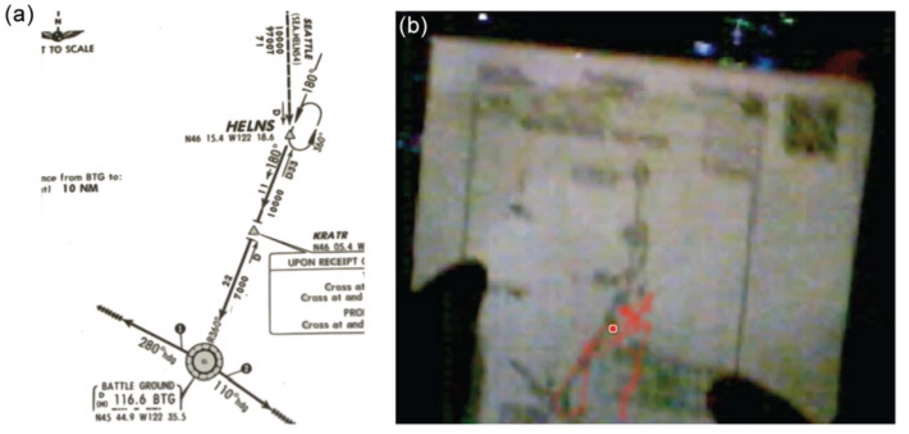
It seems plausible that as workplaces become more automated, the role of the body of the operator will diminish. However, this example also illustrates how interactions with advanced display technologies and automated systems continue to rely on the resources of the body. This example also presents our first example of a three-way interaction between two members of the flight crew and an interactive (even though sometimes unresponsive) automated system.

### Modifying the Flight Route

Ensuring that the airplane follows the planned flight path is the responsibility of the PF and in an automated airplane is primarily accomplished via visual attention. However, the eye gaze data of the PM show that he also monitored the flight path by directing his gaze to the altitude, vertical speed, and airspeed indicators on his PFD. Additionally, modifying the flight route while in flight is the responsibility of the PM. However, we observed that the PF also contributed visual and verbal attention to this task.

Because the airplane is being flown by the autopilot and autothrust systems, which are following a three-dimensional path programmed into the flight management computer system, the crew does not need to attend continuously to the control of the airplane's trajectory. This example is a clear demonstration of the idea that automation frees up attentional resources for other activities. Integrating dual eye tracking data with the audio and video recordings of both pilots in ChronoViz permits us to describe the allocation of visual attention by the flight crew as a system.

Let us look at this with a more detailed example. About 10 min into the flight, the PM was simultaneously handling flight charts, directing




(c)

[29:55.06]

PM: MK so after Helens,

Then we're gonna go to

Battle Ground.

[29:59.16]

*Figure 10.* Three parallel representations of a flight route.

visual attention to a particular flight chart, and talking about the data on the chart. While holding the arrival chart, the PM said, "MK, so after Helens, then we're gonna go to Battle Ground." This is clearly a propagation of a representation of the planned route from the navigation chart to a spoken medium. The eye gaze data provide some new insights into this process. The PM's eyes first jumped around the chart making very brief (less-than-200-ms) fixations before landing on the depiction of the HELNS (Helens) waypoint for a 400-ms fixation. The PM then readjusted the position of the chart. Once the chart was stationary again, the PM fixated for 500 ms below the depiction of the HELNS waypoint (along the path of flight), then for another 500 ms farther down the route of flight. He then made a saccade to the information box for the Battle Ground VHF omnidirectional range (VOR) and fixated there for 900 ms. Finally, there was a brief saccade and fixation to the side of the VOR symbol and a large amplitude saccade off the page, to fixate on the CDU.

This trajectory of eye gaze events can be seen to enact the planned route in the medium of eye gaze. Figure 10 shows this interaction. At this point, the flight route has been represented in the flight deck system in three ways: (a) as graphics and text on the arrival chart, (b) in the pattern of PM's eye gaze over the chart, and (c) in the words of the PM (Figures 10a, 10b, and 10c, respectively). The PM's eye gaze and speech reveal an intrapersonal dynamic configuration of visual and verbal attention in the inspection of the chart and the construction of the utterance. The representation of the route in speech lags the enactment of the representation of the route in eye gaze. This lag is expected because the pilot must know the identities of the waypoints before he can verbalize them. What is not so obvious is the way the depiction of the route on the chart, in concert with the cultural practices of chart reading in the professional pilot community, provides the resources for the pilot to enact the route in motor activity (eye gaze) by reading the chart in a particular way.

The dual eye tracking data integrated into ChronoViz also permits us to describe the allocation of visual attention by the flight crew as a system. We use the video record to navigate to the moment of interest and then inspect the eye tracking displays for both pilots, which ChronoViz maintains synchronized with all other data streams. While the PM was reading the chart, the PF was directing visual attention to the display where the PM will enter the navigation information into the flight management computer. In particular, the PF looked at the blank space on the “RTE 1 LEGS” page (displayed on the CDU) where the PM would enter the name of the next waypoint after HELNS. The PM’s spoken utterance thus coordinated the allocation of visual attention by the two pilots to two different representations of the flight route, one on the chart and the other in the waypoint list on the route legs page. This allocation illustrates a multiparty, multimodal attention configuration. While the PM’s visual attention was on the source of navigation information (chart), and the PF’s visual attention was already on the destination for that information (CDU).

This allocation of attention to the next locus of action in the ongoing navigation activity is evidence of the pilots’ joint participation in, and construction of, a shared problem-solving activity. The PF’s eye gaze anticipates the PM’s next action in the activity, which is entering the identifier for the waypoint called BTG into the list of waypoints that define the route. Further evidence of anticipation and participation in a jointly constructed activity comes from the PF’s next utterance, which he produced as the PM repositioned his body and hands to touch the keypad and enter the waypoint identifier. The PF spoke the three letters (BTG) that compose the identifier for the Battle Ground waypoint. These are the letters that the PM subsequently entered into the CDU.

## Visualizing the Dynamics of Joint Eye Gaze

The foregoing description provides details of the coordination of eye gaze for a brief interaction between the PF, PM, and airplane automation system. It is also possible to visualize the relationship of the eye gaze of the two pilots over longer spans of a flight. We use a gaze cross-recurrence plot to show this relationship. Figure 11 shows the gaze cross-recurrence plot for the flight from the performance of the preflight checklist up to and including the navigation problem solving described in the previous section.

To construct a cross-recurrence plot, we first define a set of AOIs in the region of the visual field where we have eye gaze data. The regions we chose are highlighted in colored boxes on the image of the B787 instrument panel, as shown in Figure 12.

The eye gaze data for the two pilots is first mapped into the shared coordinate space of the instrument panel. To create the recurrence plot, we consider each fixation on an AOI made by the PF. We then find all instances when the PM fixated on the same AOI. For each such match, we plot a region in the recurrence plot located at the intersection of the temporal bounds of the PF fixation and the PM fixation and colored with the plotting color associated with the AOI, as shown in Figure 12. All colored regions indicate AOIs attended by both pilots. Colored regions on the diagonal indicate AOIs attended simultaneously by the two pilots. Regions to the upper right of the diagonal indicate AOIs attended by the PM before the PF. Regions to the lower left of the diagonal indicate AOIs attended by the PF before the PM.

The event begins with both pilots attending to the right side of the PM ND. This region of the instrument panel is where the electronic checklist was displayed. These boxes indicate the visual attention allocated by both pilots to the joint performance of the preflight checklist. Following the diagonal down to the right, the large magenta box indicates the application of takeoff thrust. The PF begins attending to the engine instruments before the PM does. He also ceases attending to the engine instruments before the PM does, as is prescribed by the takeoff procedure. The next green region along the diagonal indicates the crew solving a navigation problem at that time. The PM’s ND was configured in large map mode so that the region that at other times displayed an electronic checklist or the CDU was, at that time, part of the ND. The next red region indicates the crew again reasoning

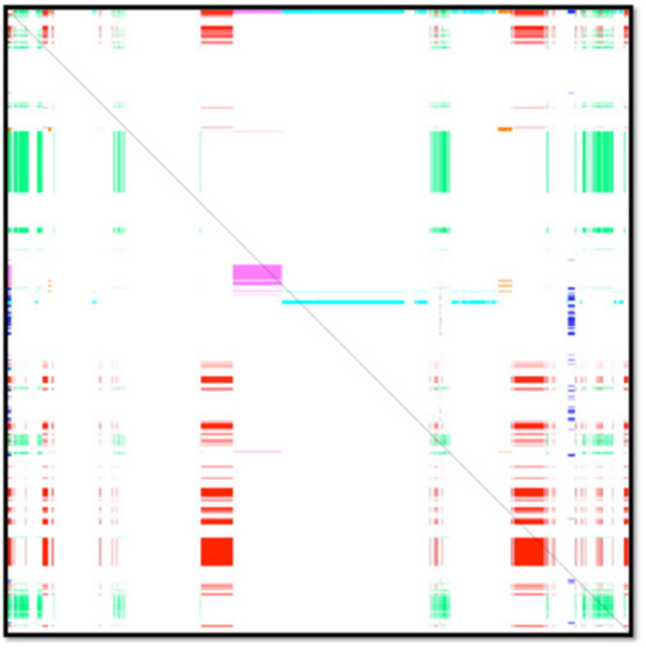


*Figure 11.* Cross-recurrence plot for 12 min of eye gaze in the Boeing 787 Dreamliner. The areas of interest (AOIs) attended by the pilot flying (PF) are shown across the top of the plot from left to right, and the ones attended by the pilot monitoring (PM) are shown down the left side of the plot from top to bottom. AOI blocks on the diagonal were viewed simultaneously by PF and PM. Time increases from left to right, from top to bottom, and down the diagonal. Gaze fixations on AOIs are color coded according to the scheme shown in Figure 12.

*Figure 12.* The seven areas of interest (AOIs) defined on the Boeing 787 Dreamliner instrument panel.

about the navigation problem jointly while looking at the PM's ND.

The final green region on the diagonal shows where the PM split his ND to display the CDU in the main instrument panel. On that display, the crew then jointly attempted, unsuccessfully, to locate and activate an approach to the destination airport. Both pilots abandoned this activity at the same moment, and the PM switched back to normal ND. This action is interesting because both pilots' eye gaze was entrained by the same display change. The entrainment gives rise to a square on the diagonal of the cross-recurrence plot. This result suggests that the recurrence plot may be an excellent way to visualize the effects or effectiveness of cues that recruit joint visual attention. In some cases, the recruitment of visual attention of both pilots is desired or intended, as when a crew alerting system message is displayed, for example. In other cases, it is appropriate for the flight crew to allocate available visual attention to different locations. When the crew encounters a collision avoidance warning while flying in visual meteorological conditions, for example, it is not appropriate for both pilots to shift their gaze to the traffic display inside the flight deck. At least one pilot should be looking outside the flight deck, conforming to the "see-and-avoid" principle for visual flight conditions.

Modifying the flight route is a key automation management activity in all modern civil transport airplanes. Routes are specified to the flight management computer as strings of characters that specify sequences of geographic waypoints with associated altitudes and speeds. To build a route leg, pilots use the CDU and construct a sequence of waypoint designations and (if required) altitudes and speeds. One interesting aspect of this activity is that the pilots are required to reconcile two very different representations of the path of the airplane. Charts and navigation displays show the route graphically as points in space connected by route legs. The representation of the route on the CDU is a list of strings of characters, whereby the only aspect of spatial layout that matters is that strings of characters higher on the page designate waypoints to be visited before those displayed lower on the page. We did not expect to see a pilot enact the route of flight in the interaction of eye gaze with an arrival chart as the PM did, but once we saw that, it made sense. It was also not surprising to see a pilot's eye gaze enact the sequential spatial relations of route waypoint designations on the CDU route legs page as the PF did. After all, reading a list from top to bottom is a highly overlearned skill.

However, the coordination of these two kinds of eye gaze enactment of route legs in a joint navigation problem-solving activity was a surprise. The structure of problem solving and look-ahead is visible in the allocation of visual attention by the flight crew system but not in the gaze data for either pilot alone. This finding means that the representation of the problem space in the flight crew system is not contained entirely inside either pilot as an isolated cognitive system.

The eye gaze cross-recurrence plot is generated by a computer program. It does require a specification of the AOIs, but once the AOIs have been defined, no frame-by-frame hand-coding of data is required. That is, the same analysis can be done with no additional work for similar data sets collected from any number of pairs of pilots. The gaze cross-recurrence plot provides a nice visualization of sequential aspects of the flight crew visual attention system. When one pilot's eye gaze generally precedes or leads the other, it is visible as a weight of plotted regions to one side or the other of the diagonal. In fact, by computing the weight of plotted space on lines parallel to the central diagonal on either side, one can quantify the extent of lead or lag between two subjects engaged in joint work (Nüssli, 2011). The gaze cross-recurrence plot also clearly shows the entrainment of eye gaze by salient display events. This property may make it a valuable tool in judging the effectiveness of crew alerting measures that are intended to draw the attention of both pilots.

## DISCUSSION

Our research makes contributions to human factors method and theory and has a variety of applications.

### Contributions to Method

We aspire to design and implement an integrative program of behavioral and cognitive research with relevance to real-world operations. The richness of expert human behavior in high-stakes activities, such as flying an airliner, presents a number of serious challenges to researchers. Real-world activity is typically messy, and participation in culturally elaborated practices, like flying, requires considerable expertise, which can make real-world activity opaque to outsiders. Our investigations of flight crew activity are grounded in an ongoing long-term cognitive-ethnographic study of commercial aviation operations (Holder & Hutchins, 2001; Hutchins & Klausen, 1996; Hutchins, Middleton, & Newsome, 2009; Hutchins, Nomura, & Holder, 2006; Hutchins & Palen, 1997; Hutchins, 1995b, 2000, 2007; Nomura, Hutchins, & Holder, 2006; Nomura & Hutchins, 2007). This ethnographic background allows us to interpret expert action and to ensure the ecological validity of our studies in high-fidelity flight simulators.

Culturally elaborated real-world activity also typically involves multiple operators in interaction with one another and with complex technical systems. Our object of study is therefore complex, multiparty, multimodal, and sociotechnical. Taking advantage of 70 years of development of sophisticated virtual reality environments, we instrument flight simulators and record the behavior of qualified flight crews in near-real-world situations. Modern sensor technology makes it possible to measure an unprecedented number of features of activity in such systems. However, the proliferation of measurements creates its own problems. Synchronizing and visualizing the relations among multiple data streams are difficult technical problems. Navigating rich data sets is difficult because of the sheer amount of data that must be confronted. Navigation can be facilitated by good annotations and metadata, but providing even minimal annotation in the form of a timeline of events is a daunting and expensive task.

Our application of ChronoViz to the problems of measuring, analyzing, and visualizing the behavior of airline flight crews provides solutions to many of these problems. We have an interface that supports the temporal alignment of multiple data streams through direct interaction with visualized data. Timelines of key events can be generated and displayed within minutes of the conclusion of a simulator session. The entire data set can be navigated via any of the representations of any of the data streams. A great deal remains to be done, but we believe we have taken some important first steps toward using computational methods and good interface design to break through the analysis bottleneck created by the need to hand-code complex data sets.

All of the examples we presented in this paper move beyond isolated aspects of human performance. Even the analysis of the heading change activity, which involved only one pilot, demonstrated the need to combine representations of multiple modalities of attention in order to understand the cognitive dimension of the activity. The challenges of measuring, quantifying, and visualizing the performance of the flight crew *system* is at the heart of our program. The gaze cross-recurrence plot is an example of the measurement, quantification, and visualization of the properties of this system. This plot makes visible properties of the system of interaction rather than properties of either of the pilots in isolation.

## Contributions to Theory

We believe that advances in methods for measuring and visualizing human activity like those described earlier can enter into a virtuous cycle in which new measures make new phenomena visible, and new phenomena challenge existing theoretical categories, which leads us in turn to look in new places and measure new things.

Among the interesting phenomena revealed by our new methods are as follows:

- The coordination of PM verbal behavior and PF elevator inputs on takeoff: The distribution of cognitive and physical effort across the two crew members in the takeoff is a beautiful example of a *distributed cognition* system. The example makes it clear that the cognitive properties of the flight deck cannot be explained in terms of the behavior of the individual pilots. The correct unit of analysis for this cognitive system is the *flight crew system*.
- Economies in the allocation of visual attention in fine hand–eye coordination made possible by the presence of the distinguishable haptic affordances of the Heading Select knob: The microscopic examination of the relations between visual and proprioceptive/haptic attention revealed an unexpected relationship between hand and eye. Now that it has been observed, it is not surprising. It should fit well with existing theories of multimodal information integration. This observation has design implications that, as far as we know, are not being considered at the moment.
- The mutual elaboration of the meanings of the elements of multimodal acts of meaning making seen in the display reconfiguration example (Hutchins & Nomura, 2011): In a disembodied theory of cognitive performance, the locations of the participants with respect to one another and their sensory access to each other's behavior would not matter. This example shows that understanding the flight deck cognitive system requires an embodied cognition approach. This example shows, perhaps counterintuitively, that increasing automation does not eliminate the role of the operator's body in expert activity.
- The enactment of the flight route in two different representational formats by the simultaneous eye gaze of the two pilots: This is a second surprising finding revealed by examining the place of eye gaze in the interactions of crew with automated systems. The interactions of enacted representations in multiple sensory and motor modalities

is just now becoming a topic of study (Stewart, Gappene, & Di Paolo, 2010) and is not yet well understood.

We are continuing to extend and expand our theoretical understanding of situated and distributed cognition. Our methods helped us identify these new phenomena that should be explained by cognitive theory. In some cases, existing theories are adequate to explain the phenomena, although reframing the phenomena in an appropriate theoretical framework makes it easier to see how existing theories apply. In other cases, existing theories do not yet capture the phenomena of interest.

## Applications

The system described in this paper can contribute to a number of important aspects of human factors practice including design, operations, and guidance for automation policy makers.

*Design.* We make two kinds of contributions to design. First, by making visible the processes that underlie performance, we provide a framework for the conceptualization of new design interventions. Second, we provide tools to measure and visualize the cognitive consequences of design interventions. Since these contributions are described throughout the paper, we will not recapitulate the all of those details here. Let us point to just two possible directions. First, taking the embodied nature of cognition seriously allows us to see more clearly the role of the body in flight deck activity, even as the flight deck becomes more automated. We predict that design changes that interfere with the contributions that are currently made by the resources of pilots' bodies may degrade performance. Second, many aspects of the allocation of attention in expert real-world performance are not consciously represented by pilots. The ability to measure and analyze these processes creates the possibility of designing work environments where flows of activity sculpt or shape the unconscious allocation of attention (interesting work in this area is being done by Bailey in a paradigm he calls "subtle gaze direction"; Bailey, McNamara, Sudarsanam, & Grimm, 2009).

*New approaches to training.* We believe that at present, few airlines are getting full value from their very expensive flight simulators. Simulators should not be used just to provide pilots a setting to practice flying activities. They should also provide a context for reflection on and critical review of performance, which could provide pilots with a better understanding of the standards they should meet and a better understanding of their own behavior. At present, however, the costs of accessing recordings of pilot activity in the simulator are prohibitive. Our system's automatic timeline generation capability eliminates the cost of annotating the recordings, and the display of the annotated timeline reduces the cost of navigating to desired events by orders of magnitude. We expect that this quantitative reduction in the cost of accessing the desired information may lead to a qualitative change in the work practices of flight instructors and student pilots.

Our tools can also provide additional measures to help instructors make judgments concerning the readiness of a pilot for qualification. We have already demonstrated the automatic measurement and visualization of student performance with respect to the so-called technical skills (the stick and rudder skills mandated by the Practical Test Standards). At present, the non-technical "competencies," such as decision making, resource management, time management, communication, and so on, are judged only subjectively. The assessment of these competencies is becoming especially important in light of the efforts to accelerate the transition from novice pilot to airline first officer as seen in the Multi-Crew Pilot Licensing initiative of the International Civil Aviation Organization. Our capabilities to measure, quantify, analyze, and visualize multimodal behavior may provide the means to automatically assess some aspects of these competencies as well. Our goal is not to replace an instructor but to provide the instructor with additional tools that will make the job of judging pilot performance easier and more accurate.

We believe our analysis techniques can be used to study crew interaction patterns in ways that inform the development of new crew resource management training practices. The ability to capture the dynamics of attention allocation and attention neglect in specific activity contexts, for example, is an excellent foundation for targeted training of pilot monitoring skills.

Are there common patterns in the details of pilot behavior that underlie superior pilot performance? Are there other common patterns that give rise to inferior pilot performance? Many flight instructors believe that such patterns exist, but at present, we lack the quantitative measures required to answer these basic questions. Because our approach enables us to document the fine-scale details of pilot performance, we believe it is now possible to document a baseline of what pilots do when they are performing well. If we could specify crew interaction patterns that contribute to effective pilot performance, we might *introduce* those patterns into training as explicit representations for pilots to apply in their daily activities. Alternatively, we might structure training activities in ways that reliably *induce* those patterns in pilot behavior whether or not the pilots are aware of the patterns.

*Automation policy.* We believe that our analyses can provide new information to consider in the trade-offs that must inevitably be balanced in the creation of regulatory guidance. For example, because crew error is implicated in a substantial fraction of incidents involving automated systems, ways must be found to reduce the incidence of crew errors in the management of automated systems. Considerable uncertainty remains concerning how best to accomplish this. Federal Aviation Regulations (FAR) Part 25 provides regulatory guidance for the manufacture of transport category airplanes. It mandates that flight guidance system functions, controls, indications, and alerts must be designed to "minimize flight crew errors and confusion concerning the behavior and operation of the flight guidance system" (FAR 25.1329[i]; http://rgl.faa.gov/Regulatory and Guidance Library/rgFar.nsf/FARSBySectLookup/25.1329). Unfortunately, no one knows exactly how to minimize flight crew errors and confusion, partly because no one knows what processes give rise to flight crew errors and confusion with respect to automated systems. We believe that the contributions to theory described earlier suggest that our approach can enhance our understanding of the processes underlying the organization of flight crew activity in interaction with automated systems.

The field of aviation human factors has been addressing the possibilities and problems of aircraft automation for many decades and has made a great deal of progress. The future of aviation-automation human factors appears bright. Recent developments in measurement technology and in cognitive theory are creating new opportunities to advance our understanding of flight crew activities.

## ACKNOWLEDGMENTS

We thank Ram Dixit, Sara Kimmich, and Alex Fung for help with preparation of figures and copyediting. Support for this research was provided by National Science Foundation Award #0729013 and UCSD-Boeing Agreement 2011-012.

## REFERENCES

Bailey, R., McNamara, A., Sudarsanam, N., & Grimm, C. (2009). Subtle gaze direction. *ACM Transactions on Graphics*, *28*, 100.

Clark, A. (2003). *Natural-born cyborgs: Minds, technologies and the future of human intelligence*. Oxford, UK; Oxford University Press.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford, UK: Oxford University Press.

Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Belknap Press.

Fouse, A., Weibel, N., Hutchins, E., & Hollan, J. D. (2011). ChronoViz: A System for Supporting Navigation of Time-coded Data. In *Extended abstracts of CHI 2011, ACM Conference on Human Factors in Computing Systems* (pp. 299–304). New York, NY: ACM Press.

Gibbs, R. (2006). *Embodiment and cognitive science*. Cambridge, UK: Cambridge University Press.

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*, 606–633.

Hagedorn, J., Hailpern, J., & Karahalios, K. G. (2008). VCode and VData: Illustrating a new framework for supporting the video annotation workflow. In *Proceedings of AVI 2008* (pp. 317–321). New York, NY: ACM Press.

Holder, B., & Hutchins, E. (2001). What pilots learn about autoflight while flying on the line. In *Proceedings of the 11th International Symposium on Aviation Psychology* (pp. 526–531). Dayton, OH: Wright State University.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer–Human Interaction*, *7*, 174–196.

Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science*, *19*, 265–288.

Hutchins, E. (2000). The cognitive consequences of patterns of information flow. *Intellectica*, *1*(30), 53–74.

Hutchins, E. (2007). Measuring change in pilots' conceptual understanding of autoflight. In *Proceedings of the 14th International Symposium on Aviation Psychology* (pp. 281–286). Dayton, OH: Wright State University.

Hutchins, E., & Klausen, T. (1996). Distributed cognition in an airline cockpit. In Y. Engeström & D. Middleton (Eds.), *Cognition and communication at work* (pp. 15–34). New York, NY: Cambridge University Press.

Hutchins, E., Middleton, C., & Newsome, W. (2009). Conceptualizing spatial relations in flight training. In *Proceedings of the 15th International Symposium on Aviation Psychology* (pp. 384–389). Dayton, OH: Wright State University.

Hutchins, E., & Nomura, S. (2011). Collaborative construction of multimodal utterances. *Embodied interaction: Language and body in the material world* (pp. 29–43). New York, NY: Cambridge University Press.

Hutchins, E., Nomura, S., & Holder, B. (2006). The ecology of language practices in worldwide airline flight deck operations. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society (CogSci 2006)* (pp. 363–368). New York, NY: Psychology Press.

Hutchins, E., & Palen, L. (1997). Constructing meaning from space. In L. B. Resnick, R. Saljo, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools, and reasoning: Essays on situated cognition* (pp. 23–40). Heidelberg, Germany: Springer-Verlag.

McNeill, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.

Nardi, B. (1996). *Context and consciousness: Activity theory and human–computer interaction*. Cambridge, MA: MIT Press.

Nomura, S., & Hutchins, E. (2007). the multimodal production of common ground understandings in intercultural flight training. In *Proceedings of the 14th International Symposium on Aviation Psychology (ISAP 2007)* (pp. 475–480). Dayton, OH: Wright State University.

Nomura, S., Hutchins, E., & Holder, B. E. (2006). The uses of paper in commercial airline flight operations. In *Proceedings of CSCW '06, International Conference on Computer Supported Cooperative Work* (pp. 249–258). New York, NY: ACM Press.

Norman, D. A. (1993). *Things that make us smart*. Boston, MA: Addison-Wesley.

Nüssli, M. A. (2011). *Dual eye-tracking methods for the study of remote collaborative problem solving*. Lausanne, Switzerland: École Polytechnique Fédérale de Lausanne.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions* (pp. 47–87). Cambridge, UK: Cambridge University Press.

Pea, R., Mills, M., Rosen, J., Dauber, K., Effeslberg, W., & Hoffert, E. (2004). The DIVER Project: Interactive digital video repurposing. *IEEE Multimedia*, *11*, 54–61.

Stewart, J., Gappene, O., & Di Paolo, E. A. (2010). *Enaction: Towards a new paradigm in cognitive science*. Cambridge, MA: MIT Press.

Suchman, L. (1987). *Plans and situated actions: The problem of human–machine communication*. Cambridge, UK: Cambridge University Press.

Weibel, N., Fouse, A., Emmenegger, C., Friedman, W., Hutchins, E., & Hollan, J. D. (2012). Digital pen and paper practices in observational research. In *Proceedings of CHI 2012, ACM Conference on Human Factors in Computing Systems* (pp. 1331–1340). New York, NY: ACM Press.

Weibel, N., Fouse, A., Hutchins, E., & Hollan, J. D. (2011, February). *Supporting an integrated paper-digital workflow for observational research*. Paper presented at IUI 2011, International Conference on Intelligent User Interfaces, Palo Alto, CA.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of LREC 2006* [CD-ROM].

Zimmerman, P. H., Bolhuis, J. E., Willemsen, A., Meyer, E. S., & Noldus, L. P. J. J. (2009). The Observer XT: A tool for the integration of and synchronization of multimodal signals. *Behavioral Research Methods*, *41*, 731–735.

Edwin Hutchins is a professor of cognitive science at the University of California, San Diego. He has devoted his career to studying cognition in real-world settings. For the past 25 years, he has been investigating cognition in the flight decks of commercial airliners worldwide. He is known for developing the distributed cognition theory and the cognitive ethnography method. He received a MacArthur Fellowship in 1985 and was elected a fellow of the Cognitive Science Society in 2002. He is the author of two books: *Culture and Inference* and *Cognition in the Wild*.

Nadir Weibel is a research scientist at the University of California, San Diego, where he investigates how to apply multimodal technology to diverse aspects of cognition and human–computer interaction (HCI). He is a member of the DCog-HCI Laboratory and the Ubiquitous Computing and Social Dynamics research group. His research is situated at the intersection of computer science, design, and cognitive science. He is developing tools, techniques, and infrastructure to support innovative interactive devices and is studying the cognitive consequences of their introduction in real-world contexts.

Colleen Emmenegger is a staff research associate in the DCog-HCI laboratory at the University of California, San Diego, where she supervises the collection and analysis of multimodal data on the activities of pilots. Her interests lie in understanding distributed cognition in complex and "messy" real-world settings and exploiting tools and techniques that support research domains rich in data. She also enjoys working with, and learning from, bright and innovative student interns.

Adam Fouse received his PhD from the Department of Cognitive Science at the University of California, San Diego. His research focuses on interactive information visualization of time-based activity to aid analysis of observational data. Using methods from those of cognitive psychology, cognitive ethnogra-

phy, and computer science, he combines an understanding of real-world use of visualization technology with development of new tools, most notably, the ChronoViz data visualization and analysis tool.

Barbara Holder is lead scientist and associated technical fellow at Boeing, where she leads the Flight Deck Concept Center, a multimillion-dollar research and innovation center devoted to investigating flight deck design and operations concepts. She conducts field research at airlines around the world to investigate the role of culture and technology in the conduct of pilots' daily work. Her research has been applied in Boeing flight decks, crew training and procedures, checklist design, and operational policy.