HCI, Natural Science and Design: A Framework for Triangulation Across Disciplines

Wendy E. Mackay & Anne-Laure Fayard

Centre d'Études de la Navigation Aérienne
Orly Sud 205
94542 ORLY AÉROGARES
FRANCE
mackay@lri.fr *fayard@cena.dgac.fr*

ABSTRACT

Human-computer interaction is multidisciplinary, drawing paradigms and techniques from both the natural sciences and the design disciplines. HCI cannot be considered a pure natural science because it studies the interaction between people and artificially-created artifacts, rather than naturally-occurring phenomena, which violates several basic assumptions of natural science. Similarly, HCI cannot be considered a pure design discipline because it strives to independently verify design decisions and processes, and borrows many values from scientists.

The purpose of this paper is to provide a simple framework that describes how the research and design models underlying HCI can be integrated. We explore the relationships among these approaches in the context of a particular research site, CENA, the Centre d' Études de la Navigation Aérienne, and illustrate how the various disciplines can contribute to a complex design problem: improving the interface to the French air traffic control system.

The framework provides one perspective for understanding the various research approaches, and, more importantly, suggests new research directions. The resulting cross-disciplinary triangulation can increase the effectiveness of the individual research and design approaches.

KEYWORDS: CSCW, design, theory, augmented reality

INTRODUCTION

How do we design new computer systems that support expert users in complex, real-time collaborative work environments, especially when they already have well-established, successful work practices and no tolerance for error? The field of Human-Computer Interaction addresses such questions with a multidisciplinary approach, drawing tools, techniques and paradigms from a variety of existing disciplines in both science and design.[1] Members of the HCI community pursue a variety of goals ranging from providing general theories and principles, to reporting on detailed observations of actual use, to creating innovative new designs.

How do we decide among the multitude of paradigms available to us? Unlike researchers or designers working within a single academic discipline, with well-established procedures for conducting their work, we find ourselves constantly borrowing, inventing and re-inventing techniques as we go. We draw from both science and design and must be able to converse with researchers and designers completely immersed in their individual disciplines. We work at both applied and theoretical levels (and find most of the debates challenging whether a particular study is "scientific enough" to be unproductive). Because we create working prototypes, we select the methods that seem most appropriate for the problem at hand. At the same time, we must conduct our work in a way that is fundamentally sound at the level of each discipline we draw from and viewed as legitimate by our academic colleagues.

The purpose of this paper is to provide a framework for viewing the different disciplines that contribute to HCI, including the epistemological origins of the scientific branches. Our hope is to draw attention to the most important assumptions underlying each and to point out their key similarities, especially when they are based upon the same underlying scientific approach. We also make a fundamental distinction between the sciences and design and show why a multidisciplinary field like HCI must necessarily draw from and benefit from both. Rather than discuss these at a purely theoretical level, we use our own research setting to illustrate how a wide variety of

[1] Our own backgrounds include a mix of disciplines: experimental psychology, study of innovation & software product design, and philosophy & cognitive science.

DIS '97 Amsterdam, The Netherlands
© 1997 ACM 0-89791-863-0/97/0008...$3.50

approaches can be useful in addressing a complex user interface design problem.

The paper first introduces the research environment at the CENA, which conducts research on the French air traffic control system. We then present a model that illustrates how HCI draws from both the natural sciences and design disciplines and discuss key assumptions of each. We then discuss why HCI is neither a science nor a design discipline, because the focus of the work is the study of the interaction between people and artificially-constructed artifacts. We use the range of research strategies used at CENA to illustrate how science and design disciplines can be integrated to address a complex, real-world HCI problem. We conclude with a general discussion of how designing interactive systems can be effectively served by techniques drawn from science, design and new techniques constructed specifically for HCI. The framework presented in the paper can help suggest new research directions and facilitate triangulation, i.e. the use of different methods to address the same problem.

RESEARCH SETTING: CENA

CENA, the Centre d'Études de la Navigation Aérienne, is responsible for the design and implementation of new technologies to support the air traffic control system in France. The Paris center for *en route* air traffic control, located in Athis Mons next to Orly airport, handles one of the most complex air traffic patterns in Europe. This is due both to geography, given the importance of the London-Paris-Frankfurt triangle, and history, given the difficulty of flying over Eastern Europe during the cold war.

Like its counterparts around the world, the current system was designed and introduced after World War II and the use of RADAR and paper flight strips has not changed fundamentally since the 1960s. The French system is extraordinarily safe: since its inception in the early 1950's, no plane crashes have been attributed to the civilian air traffic controllers.[2] Of course, the system has evolved over the years: improved RADAR and other technologies have been introduced and the organization of the flight sectors have been changed to meet changing traffic patterns. Even so, the basic methods remain the same, with relatively minor variations evident at the country, flight center, sector, team and individual level.

If the system is so safe, why change it? The first reason is technical: much of the existing technology is very old and difficult or impossible to replace. If the systems must be replaced anyway, it makes sense to take advantage of the tremendous advances in computing achieved over the past few decades. The second reason is both ethical and political: levels of air traffic are increasing rapidly and the current system is extremely, but not completely, safe. A recent article in the International Herald Tribune (15 Jan. 1997) warns that if current safety levels are maintained and the traffic continues to increase at its present rate, a plane will crash somewhere in the world every week by the year 2004. Airline management and government officials are concerned that the public will perceive this as unacceptably high (despite the fact that flying remains several orders of magnitude safer than driving a car if we look at kilometers traveled). Either the level of safety must be increased or limits must be placed on the number of flights. Given the economic consequences of the latter, it is not surprising that attempts are underway to reexamine and improve the current air traffic system.

CENA is responsible for research in France on air traffic control. with researchers from a variety of disciplines, including engineers, computer scientists, cognitive ergonomists, ethnographers and former air traffic controllers. Clearly, all of these skills are required, given the complexity and difficulty of the task at hand. Yet, given the difficulties of communication across disciplines, much of the work tends to be done in isolation. The phenomenon is familiar in the HCI community: system designers often have little interaction with controllers or the researchers who study them. Although controllers are often asked their opinions or invited to test simulations of the system, they are rarely active throughout the design process.

A major project has been underway for over a decade that has resulted in the purchase of large, high-definition computer screens and the development of a major software system to replace the existing one. The "furniture" designed for the system is about to be introduced, but questions are now being raised about some of the software design decisions. Attempts to create electronic versions of flight strips, accessed by a mouse and a computer keyboard, appear to be slower and more prone to error than writing on paper flight strips. As in similar attempts to automate highly complex collaborative work settings, e.g., Suchman, (1987), Zuboff (1988), Barley et al. (1988) and Mackay (1990), some important, possibly invisible, aspects of the current system do not appear to have an effective counterpart in the new system. The problem now is to find an effective solution: we need to build upon existing results and techniques and develop a prototype system that takes advantage of modern computer technology, increases safety without loss of efficiency and respects and enhances the existing work practices of air traffic controllers.

Ethnographic studies can illuminate important aspects of the work setting that must be accounted for in any new system. Heath and Luff (1991), in their extensive studies of the management of subway trains in the London Underground, highlighted the importance of the peripheral awareness of each other's activities. What appears to be a single-user task is in fact highly dependent upon an implicit understanding of the work of others. Studies of air traffic controllers, e.g., Gras et al., (1994) and Hopkin (1995)

[2] A mid-air collision occurred in 1973, when military air traffic controllers took over temporarily during a civilian controller strike. The strike was quickly resolved and French controllers have had significantly more power and better working conditions since then.

emphasize the complexity and subtlety of the interactions among controllers. Preux (1994) studied the specific role of the paper flight strip, not only as an individual memory aid, but also as a focal point for cooperative work. Studies of the English en route air traffic controllers, Hughes et al. (1992) and Bentley et al. (1992) confirm the critical role of the flight strip and question the wisdom of replacing them with electronic imitations that reduce the level of cooperative work. Harper et al., (1991) concluded that it is the gestalt of the strips, rather than the absolute information they contain, that facilitates cooperation among team members and hand-offs between shifts.

Our own research is based upon an ethnographic study in which we followed a team of controllers through their schedule for a period of four months. We were particularly interested in the role of paper flight strips and our observations match those described in the English system (with some interesting cultural differences.) The next step is to build a series of design prototypes. We have chosen to pursue a different design direction than in the English studies and are exploring an approach called "Augmented Reality" (Wellner, Mackay & Gold, 1993, Mackay, 1996). The goal is to preserve and augment rather than replace existing real-world artifacts. In our specific case, we are trying to create a system that preserves the flexibility and subtlety of the paper flight strips, while capturing information that will be useful in tools that aid the controllers.

To be successful, we must draw upon techniques from a variety of research and design disciplines. We must understand as much as we can about the actual activities of controllers and their current interaction with paper flight strips, in both qualitative and quantitative terms. We must generate new design ideas, using researchers, designers and users (controllers) to provide innovations. Finally, we must be able to distinguish among different design possibilities and select those that make the most sense in the current context.

The techniques require the full range of disciplines found in HCI. The next section presents a framework that describes them and discusses their underlying assumptions, placing them within a unified context for design.

NATURAL SCIENCE, DESIGN AND HCI

Human-computer interaction, like other multidisciplinary fields, borrows techniques from its component disciplines and must determine how they relate to each other. While it may be appropriate for some to engage in heated debates about competing approaches, it is rare to find a designer who finds this useful. What is more important is to understand enough about each approach and the corresponding assumptions to be able to choose which is most appropriate for addressing a particular design problem.

Figure 1 presents some of the component disciplines that contribute to the study of human computer interaction. The full range of natural and social sciences is represented by the box at the top. Although they differ greatly in subject matter and specific methods, they all derive their basic values and assumptions from a general model that governs all scientific disciplines. The corresponding box at the bottom represents schools of Fine Arts and the engineering and design disciplines. These also differ greatly in subject matter and specific methods, but share the goal of creating new artifacts.

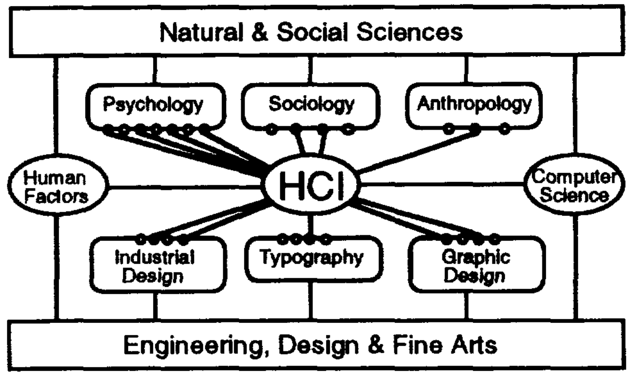

Figure 1: Human-computer interaction, like human factors and computer science, is a multidisciplinary field that draws from both scientific and design disciplines.

The sciences are divided into separate disciplines, such as Psychology, Sociology and Anthropology, represented by rounded rectangles. Scientists within these disciplines operate within *paradigms* (Kuhn, 1962, Latour, 1987), that dictate which questions are considered "interesting" and which specific methods can be used to address those questions. Within any branch or sub-area, researchers can range from highly theoretical to highly applied. Various engineering and design disciplines, including Graphic Design, Typography and Industrial Design are also represented by rounded rectangles. Designers and artists in particular operate within schools that dictate aesthetics and "style".

The ovals in the center represent multidisciplinary fields. Each multidisciplinary field draws concepts and methods from particular branches (shown as small circles) of its component scientific and design disciplines. For example, HCI borrows from Cognitive Psychology and Ethnography (branches of Psychology and Anthropology) as well as typeface design (a branch of Typography). HCI also draws from other multidisciplinary fields, particularly Human Factors and Computer Science. The latter can be considered one of the "Sciences of the Artificial" (Simon, 1969), yet still requires aspects of engineering and design.

Figure 1 is not meant to be exhaustive, but rather to illustrate why HCI is neither a scientific nor a design discipline, but actively incorporates aspects of both. Although similar in some respects, science and design differ in many assumptions and values. For example, while both scientists and designers are trained in an apprenticeship, the measures of success are quite different. A science graduate student works in a laboratory with a senior scientist; develops an original thesis and conducts original research to support that thesis, which is evaluated by other senior scientists. The terminal degree is a Ph.D. and the expected career path is to become either a professor or a research scientist in an industrial laboratory (such as Bell Labs or Xerox PARC) or a government laboratory (such as the American National Institute of Health or the French Centre National de la Recherche Scientifique). Design students usually work in design studios with other designers. Engineering and computer science students usually have internships where they work with senior engineers in industry. Students create a variety of artifacts, some assigned as learning exercises and some in the context of real projects. The terminal academic degree for a designer is usually a Master's degree, which is awarded based upon a portfolio or collection of previous work. Computer programmers similarly demonstrate their prowess based on programs created for class projects or products developed during internships.

Some design areas are more theoretical, such as cubism or impressionism in the fine arts. Others are more applied, such as typography and graphic design. Unlike scientists, who generally believe in a notion of "progress" and accumulation of knowledge, designers create new artifacts within a series of evolving styles. A few design schools emphasize the perspective of the user. For example, the notion of a pattern language (Alexander et al., 1977) encourages architects to think about the perspective of the users of the spaces being created.

HCI is a new field and, not surprisingly, most senior members of the community come from other branches of science or schools of design or engineering. Unfortunately, people trained in one discipline often find themselves using techniques from another area without understanding the underlying assumptions and techniques. This causes problems when the techniques are misapplied, most notably in the statistical analysis of experiments, but in other areas as well.

HCI and Natural Science

This section provides a brief explanation of the basic scientific approach, comparing and contrasting the two most important historical models: the deductive (Figure 2) and the inductive (Figure 3). The deductive model attempts to deduce or derive properties of the real world from theory, whereas the inductive model attempts to induce or generalize theories from observations of real world phenomena. (See Lo See (1993) for an introduction to the history of science.)

The boxes in each figure represent phenomena that exist in the real world. The scientific method provides specific techniques for moving cyclically between the theoretical level, which describes natural phenomena in general terms, and the empirical level, in which people describe phenomena they observe in the real world.

The Deductive Model dates back to Galileo and is also referred to as the hypothetico-deductive model. The purpose is to generate a set of hypotheses that can explain real-world phenomena: It is the most general model in epistemology and philosophy of science (Gillies (1993) and O'Hew (1989) provide good introductions to the philosophy of science).

The scientist begins with a theory about a particular phenomenon. She makes a specific prediction, in the form of a hypothesis, about the behavior of the phenomenon (first box, Figure 1). An experiment, usually conducted in a laboratory, systematically manipulates a set of independent variables and measures the results with respect to a set of dependent variables. Other factors that may affect the results are either eliminated or varied systematically through control conditions. Once the results of the experiment are analyzed, the original theory is re-examined and a revised hypothesis is created. The scientist can then proceed to test the revised hypothesis with a new, more precise, controlled experiments. The scientists values *reliability* , which means that the same results will be obtained if the experiment is repeated under the same conditions, and *validity*, which means that the results can be generalized beyond the specific experimental setting in the laboratory.

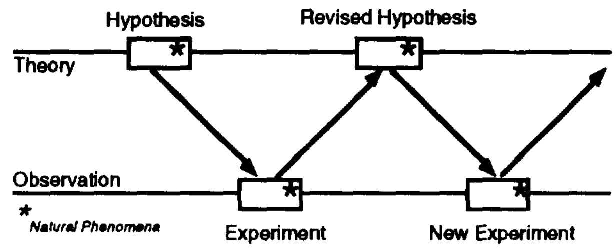

Figure 2: Deductive Model: Begin with a theory about a naturally-occurring phenomenon, generate a testable hypothesis, test it with a controlled experiment, revise the hypothesis and test again with a new controlled experiment.

Of course, theories must come from somewhere. Scientists present their theories in the context of other theories and previous observations of the real world. However, in the basic deductive method, the cycle always begins with a theory, followed by an experiment or test. Within the social sciences, Experimental Psychologists are most likely to follow this model.

The Inductive Model dates back to Bacon's conception of science and the inductive method. The purpose is to construct the best description (as opposed to explanation) of the real world. In contrast to the Deductive Model, the Inductive Model begins with natural phenomena observed in the real world (box 1, Figure 3).

The scientist observes phenomena in the real world (as opposed to the laboratory), without having a preconceived idea of what he is looking for. He then attempts to describe a framework or model that explains the phenomenon. It is assumed that questions will emerge, so the next step is to return to the real world to make further specific observations that validate or contradict the original framework. The results are then incorporated into a modified framework. Traditional Sociology and Anthropology follow this model.

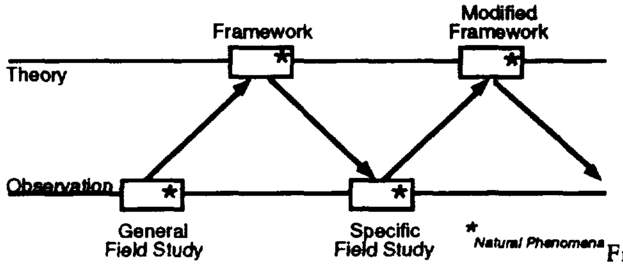

Figure 3: Inductive Model: Begin by observing phenomena in the real world (avoiding pre-conceived notions), develop a framework to describe the phenomena, make new observations with respect to questions that emerged from the original observations, and modify the framework as required.

The Deductive and Inductive Models share an important characteristic. Both assume the existence of two "worlds": the theoretical and the empirical (or "real world"). For both, research consists of moving cyclically back and forth between theory and observation in an attempt to understand phenomena in the real world. Both models may incorporate both quantitative and qualitative data, although the deductive model is more often quantitative and the inductive is more often qualitative. Standard statistical approaches, e.g. hypothesis testing and regression analysis, can be applied to the deductive model. Exploratory data analysis (Hartwig & Dearing, 1979, provide a useful introduction whereas Tukey & Mosteller, 1977 provide an in-depth presentation) is based upon the same underlying mathematical models, but is used for quantitative analysis of inductive approaches. The goal of exploratory data analysis is to iteratively identify and display the "smooth" (i.e., the general distribution(s) of the data) and the "rough" (i.e., the variability or error in the data).

Both the deductive and inductive models adhere to the scientific method and make certain assumptions:

- Natural phenomena exist and can be isolated for study.
- Observers are unbiased.[3]
- Repeated observations under the same conditions will yield the same results.
- Conclusions drawn from observations in one setting can be generalized to other settings.

The main difference between the models lies in the starting point of the cycle: The Deductive Model starts from theory and the Inductive Model starts from observation. These models also differ in the specific ways in which theories are stated and observations are conducted (Chalmers 1976).

Challenging the basic models

The scientific method has proven extraordinarily productive and most working scientists in the natural sciences follow variations of these models. However, it is important to note that these models have been severely critiqued and some of the most basic assumptions have been challenged.

Popper (1963) challenged both models, on different grounds. He disagreed with the Inductive Model's approach of starting with observation. However, he also disagreed with the Deductive Model's approach of attempting to confirm a particular hypothesis. He claimed that, although it is possible to find data to corroborate a hypothesis, one can never actually confirm it. A counter-example is always theoretically possible. His solution was to pose the hypotheses differently, so that experiments try explicitly to find a counter-example to the hypothesis. The hypothesis can be accepted as conditionally true, until an objection can be found. This strategy is called "falsification" and forms the basis for most of today's experimental designs that fit the Deductive Model. Note that this implies that experiments can never actually prove hypotheses: they can only increase our confidence in them. This is the reason for the careful wording necessary when reporting the results of statistical analyses of experiments.

The Inductive Model has also been challenged. Many anthropologists now reject the notion of "unbiased observers". Margaret Mead popularized the notion of "cultural relativism", which assumes that researchers bring their own biases to the situation. Many ethnographers, who make long-term, detailed observations of people in other cultures, now explicitly try to identify as many of their own biases as possible. They also encourage their readers to do the same and judge the work in the context of these three perspectives.

Each research strategy has strengths and weaknesses. McGrath et al. (1982) argue persuasively that individual experiments or studies all have serious flaws and cannot, by themselves, be considered "answers".

Choosing a Research Model

Why do scientists choose one model over another? Sometimes the answer is practical: the phenomena being studied greatly affect the theories and types of observations used. For example, an experimental psychologist studying the perception of color can be reasonably confident that experiments conducted in the lab will be relevant outside of the lab. This is less clear for a sociologist studying the effects of socio-economic status on political beliefs and is clearly absurd for an anthropologist interested in coming-of-age rituals in aboriginal peoples.

Another important factor is the time frame: psychologists tend to study phenomena that occur within a range of a millisecond to an hour. Sociologists tend to study phenomena that occur over hours, days, weeks or years. Anthropologists tend to study phenomena that are grounded within a long-term historical context of years, decades or even centuries. Clearly, the ways of looking at these phenomena must differ.

As a field, Human-Computer Interaction poses a wide range of questions, crossing the boundaries of these disciplines. Why don't we simply consider HCI to be a natural science and select particular methods as needed? The problem is that

[3] This is a controversial point: many social scientists argue that they are necessarily biased in their observations and attempt to make those biases explicit.

we are interested in artifacts made by people and the design process is quite different from the scientific method.

Figure 4 illustrates how designers and engineers work from early prototypes to finished products, using guidelines and rules of thumb (e.g., Smith and Mosier, 1984 and Mayhew, 1992) or principles drawn from experimental Psychology (Fleming and Levie, 1978).

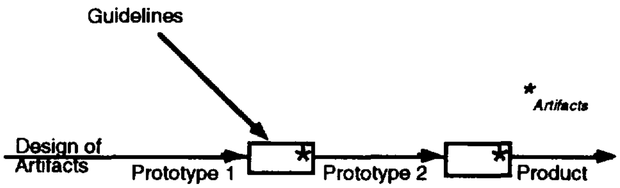

Figure 4: Designers and engineers work with guidelines and rules of thumb to create new technology.

The major difficulty for HCI is that the object of study is not an independent natural phenomenon, as in all of the sciences. Nor is it solely the creation of new artifacts, as in the design and engineering disciplines. HCI studies the *interaction* between people and artificially-created artifacts. The study of dynamic phenomena, such as human learning, is already problematic for the social sciences (Sidman 1977). HCI compounds the problem by constantly changing the underlying technology being studied. Even more difficult is the finding that the interaction of people and technology is *co-adaptive* (Mackay, 1990): people both adapt to the technology and they actively adapt it for their own purposes. Thus, the problem is not static: the "same" technology is often very different in different environments. People are influenced by how the people around them interact with the technology.

Not surprisingly, the HCI community has adopted iterative design as a way of addressing the changing nature of user's interaction with changing technology (Norman and Draper, 1986). By studying users, both to evaluate the technology and to generate design ideas, we can build better systems. How can we view these techniques with respect to the deductive and inductive models of science? Figure 5 illustrates how to integrate the various aspects of the two major scientific approaches, as well as design, to represent the range of activities in Human-Computer Interaction.

Figure 5 illustrates how design and both scientific models can be integrated for the range of activities found in HCI. At the theoretical level, we can create and revise interaction models based upon observations of users interacting with artifacts. At the empirical or real-world level, we can observe how people interact with various technologies and develop models of use. In both cases, we can draw from theory and observation to instantiate new artifacts, ranging from early simulations to working prototypes to products. These artifacts constantly evolve and influence or change models at the theoretical level and observations at the empirical level. One can argue that scientists strive to understand whereas designers strive to create. HCI is an interdisciplinary field that must do both.

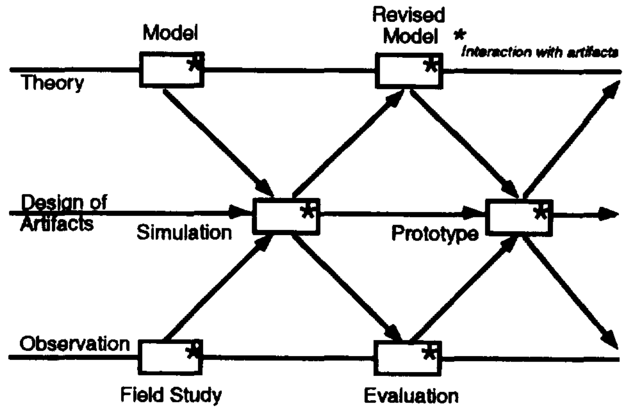

Figure 5: HCI creates and examines the interaction of people with artificially-created artifacts, moving between theory and empirical observation. The boxes represent the *interaction* with artificially-created artifacts, rather than independent natural phenomena as in Figures 2 & 3 or design artifacts as in Figure 4.

Particular design techniques can now be mapped to this perspective. Task analyses or models such as GOMS (Card et al., 1983) provide cognitive models of users' interactions with artifacts and are essentially deductive approaches. In contrast, scenario-based design (Carroll, 1995, Mackay & Bødker, 1994) and participatory design in general (Kyng & Greenbaum, 1991) draw their ideas from observations of interactions in the real world and are essentially inductive. System software evaluation techniques, such as Cognitive Walkthroughs (John & Packer, 1995) and Heuristic Evaluations (Nielsen & Phillips, 1993) begin with design artifacts and use guidelines to evaluate them whereas Design Rationale (McLean et al., 1991) seeks to preserve explanations of how the designs were created. Researchers working within the HCI paradigm shown in Figure 5 may choose any or all of the above.

Applying the Framework at CENA

CENA conducts or sponsors a wide variety of different research projects. We decided to evaluate our framework by examining whether it captures the full range of activities, both scientific and design. The following six projects are derived from different groups within the organization and represent a broad range of goals and methods. Each project can be classified initially, based upon the way the participants talk about the goals of their projects. These goals correspond roughly with the approaches outlined in Figures 2, 3 and 4, i.e., developing theories (deductive model), enriching descriptions (inductive model) or creating new technologies (design model). In each case, the basic model, whether scientific or design-oriented, provides a good indication of the focus of the project.

However, even though particular projects are grounded in one of these three approaches, they rarely use one approach exclusively. Conducting research and design activities that meet the needs of a particular user population generate specific requirements that force each project to move outside of the basic paradigm. None can be described exclusively on the basis of Figures 2, 3 or 4: yet all fit into the more complex, interdisciplinary framework presented in Figure 5.

This section briefly describes six recent or on-going projects at CENA and maps them onto our HCI framework. For each project, we identify the goals, the backgrounds of the researchers or designers, and the basic project design.

Specialists in circadian rhythms and the medical corps at the Laboratory of applied anthropology in Paris conducted a 3-year study of the sleep patterns of controllers, using interviews and physiological tests (Figure 6). An important finding was that controllers have lower fatigue and increased vigilance when their schedules begin at successively later times during the day. Although the original study did not consider the practical consequences for the controllers, another internal group responsible for work schedules was able to use the findings to revise the existing schedules. The 10-day pattern of work and rest days has been changed to 12 days, with two different kinds of night shifts (DNA, 1997). The resulting schedule is much more complex and the controllers have not yet decided whether or not they like it.

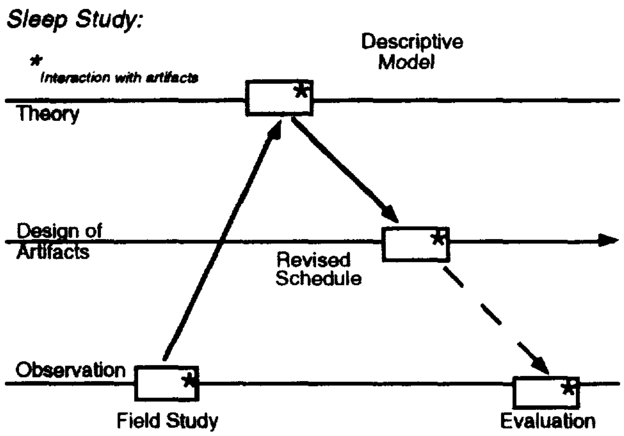

Figure 6: A study of the controllers' sleep patterns: The researchers examined the existing schedule and, together with their knowledge of human biological rhythms, created a new schedule, which was implemented 12 January, 1997.

This study combines applied biology and applied psychology. The field study approach is important to determine how the controllers' ability to interact with the control system changes according to their sleep patterns. The result of the study has a direct impact on the controller's lives, through the change in the schedule.

Figure 7 presents work done on SIM-COOP, one of a number of projects being conducted by ARAHIIMS.

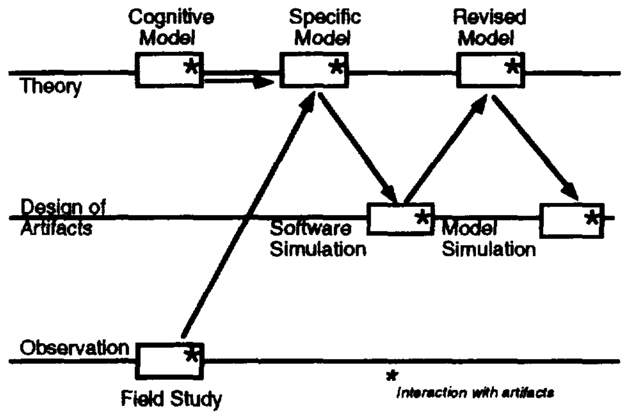

Figure 7: SIM-COOP: The ARAMIIHS group focuses on developing cognitive models of air traffic controllers, informed by observation in the field and simulated in Prolog.

ARAMIIHS is an external research group that works with CENA. As cognitive ergonomists, they are interested in analyzing the tasks and communication patterns of the controllers and building theoretical models to describe the activities. They draw from several theoretical paradigms, including speech act theory and relevance theory. Some of the work involves observation of controllers in the field, while other research involves creating cognitive models, which are then implemented in Prolog (e.g., Simcoop, (Zorola-Villarreal et al., 1996) and used to simulate a variety of air traffic control situations. The research paradigm strongly emphasizes the theoretical level, through the construction of cognitive models, but also sees the value in conducting field studies to understand specific types of interaction in the field.

Figure 8 presents CINDI, an experiment designed to compare different interaction techniques using the mouse and keyboard, under varying levels of mental load and interruptions.

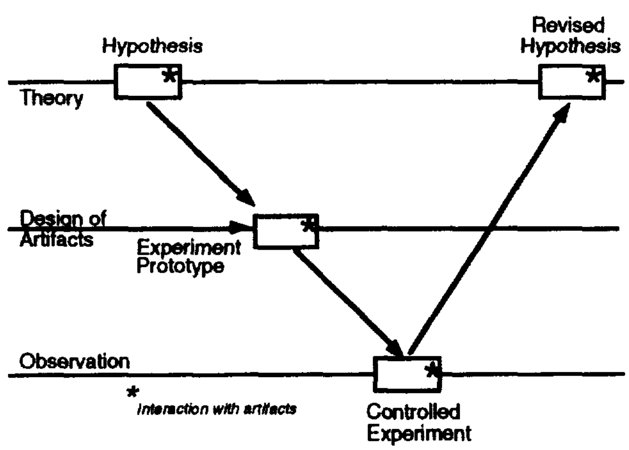

Figure 8: The CINDI project examined which of several interaction techniques was most effective for changing flight information. Subjects were systematically exposed to a variety of conditions, using an experimental prototype in a laboratory setting, and response times and errors were measured.

Air traffic controllers frequently issue commands to change flight level or direction, often while being interrupted or working on another task. Future interfaces that involve "electronic flight strips" would require controllers to issue these commands to a computer. *CINDI* (Mertz, 1996) compares a set of different interaction techniques, selecting a new flight level from one of several menus or typing it on the keyboard, in a situation designed to simulate the stress and interruptions common in real air traffic control settings.

CINDI uses an experimental protocol based on the deductive model used in Experimental Psychology. Subjects, with varying backgrounds including computer science, engineering, ergonomics and air traffic control, were asked to make a series of changes to existing flight levels. Mertz created an experimental prototype using existing 3 menus, 2 modified menus and a keyboard interface. "Mental load" was operationalized by creating an increasingly difficult set of tasks. In the simplest condition, subjects were asked to change a series of flight levels using one of the interaction techniques. In the more complex condition, subjects were asked to perform a parallel task: a red or green light would appear on the screen and the subject would press the corresponding left or right foot pedal. The most complex condition added a third parallel task: subjects were asked to memorize and later recall a set of words spoken by the experimenter.

The null hypothesis stated that no differences would obtain among the various interaction styles, regardless of the level of mental load. The dependent variables were time to completion and numbers of errors. Controlled experiments such as CINDI are rare at CENA. However, they provide an effective method of getting information about specific design alternatives. This research paradigm strongly emphasizes the deductive approach and is a classic example of a social sciences experiment.

Figure 9 presents ERATO, a system designed to support air traffic controllers as they try to manage a stream of upcoming air conflicts.

ERATO (Leroux, 1993a, 1993b) is being designed by a former air traffic controller who was originally trained as an engineer. Based on his experience in the field, Leroux developed a descriptive model of the problems controllers face when trying to resolve a series of air traffic conflicts. For example, they cannot simply solve problems as they appear, but must wait until the appropriate time to issue new instructions to the pilots. The controller might decide to ask the pilot to change direction after reaching a certain beacon, estimated to be in about 15 minutes at the current speed. The problem is to remember what to do at the right time, in a highly interrupt-driven environment. ERATO introduces the concept of an *agenda* or timeline that provides controllers with a temporal view of upcoming conflicts and reminding the controller when it is appropriate to take action.

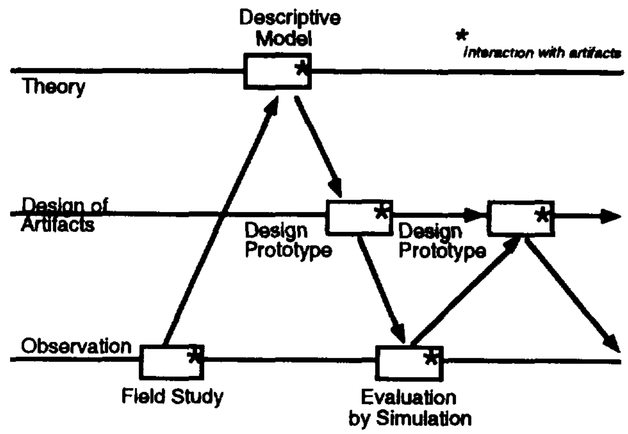

Figure 9: ERATO. Based on field observations and a descriptive model, the project is now iteratively designing a prototype that is being tested with simulated air traffic using real air traffic controllers.

Leroux is now working with an interdisciplinary team to create a working prototype with a variety of such tools. Different versions of the prototype are tested by creating simulations and inviting current air traffic controllers to try the system. The simulations are videotaped and the group looks at the differences between the controllers performance in settings with and without ERATO. The goal of ERATO is to provide an effective tool that helps controllers make informed decisions under stressful, high conflict conditions. The perspective is to provide advice rather than replace the controller. The on-going participation of a small group of currently-active controllers is considered essential to the success of the project. Although ERATO fits into the design model, it is clear that the project is also informed by both observation and a descriptive model.

Figure 10 presents GRIGRI, part of the Imagine project.

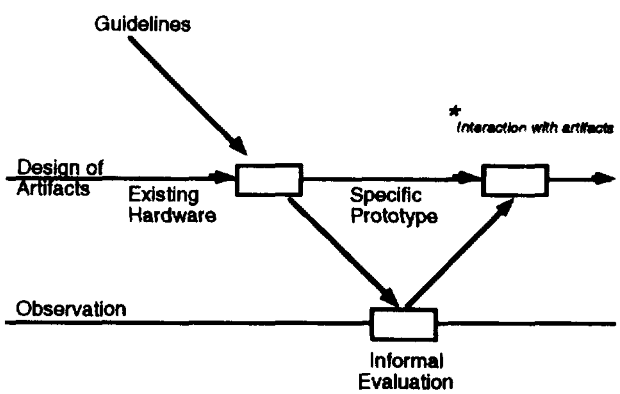

Figure 10: GRIGRI: Based on informal observations of writing on paper strips, a prototype that provides pen-based input to air traffic control was created and evaluated.

GRIGRI (Chatty & Lecoanet, 1996) also follows the design model, but from a very different perspective. The Imagine engineering group were interested in whether a particular hardware device, a graphics tablet, would be useful to controllers, given the existing activity of writing on paper flight strips. They created a set of standard marks for common activities, such as changing direction or modifying the flight level. They then displayed an electronic version of the flight strip on a screen and asked controllers to write on the graphics tablet, modifying the strips with the standard marks. GRIGRI then interprets the marks and makes the appropriate modifications to the strip.

Unlike CINDI, the GRIGRI tests were conducted with real air traffic controllers. Rather than concentrating on basic performance issues, the goal was to determine whether the idea was worth pursuing and whether or not controllers would find the approach acceptable. GRIGRI fits easily into the design model, yet provides an evaluation step that is deductive in nature.

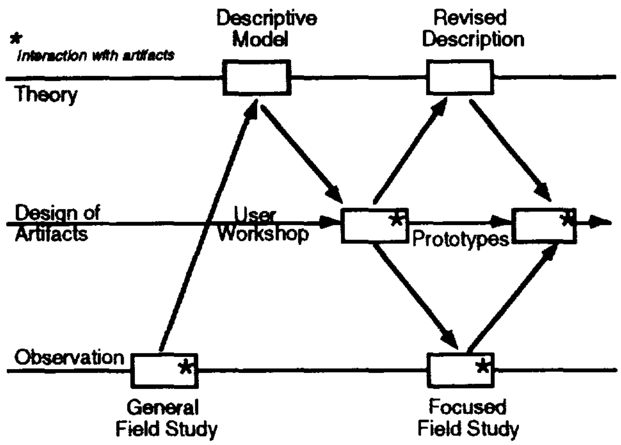

Figure 11: The augmented flight strips project (Mackay & Fayard, in preparation) is a design project with both inductive and deductive aspects.

Figure 11 presents our own work. Like the English studies of air traffic control (Hughes et al., 1992, Bentley et al., 1992), we use a combination of ethnography and design. However, our direction is different: augmented reality rather than another mouse/keyboard interface. We seek to maintain the existing paper flight strips, with all their subtlety and flexibility, and augment them, by capturing information and displaying it to the controllers. We feel that augmented reality is a very promising research direction because we ground it not in the technology itself, but in the existing work practices of the organization. Our prototyping strategies are also different: we are explicitly trying to create opportunities for controllers to generate innovative ideas.

The project began with an in-depth study of one team of controllers under a full range of traffic conditions. We are performing both qualitative and quantitative analyses of the data, which includes 40 hours of video, detailed notes and informal discussions with controllers. We are particularly interested in the cooperative interactions among controllers with respect to paper flight strips, the RADAR and other technologies.

We also use a variety of design strategies, including various rapid prototyping techniques and user-researcher workshops that generate ideas via brainstorming, scenario-building and evaluation of prototypes. Our goal is to innovate: not just to create something new for controllers, but also to explore the variety of ways in which we can augment useful artifacts in real user settings. The final result will be two or more prototypes that will be formally evaluated at the end of the project.

TRIANGULATION IN RESEARCH & DESIGN

Although the group of projects we have chosen is not exhaustive, we feel that it is representative of the range of projects conducted at CENA. The different research groups are easily distinguishable, not only by their goals but by their different approaches. Yet many aspects of the work are similar or complementary and few of the projects fit into a standard research design. How can we take advantage of this range of activities?

McGrath et al. (1982) explore the problems of conducting research in the social sciences. They argue persuasively that each individual method makes trade-offs between the fundamental problems they want to avoid and the problems they are willing to accept. Figure 12 (Source: Runkel and McGrath, 1972) shows 8 different research strategies, each of which forces a trade-off among different threats to validity (Cook & Campbell, 1979). They state that, all other things being equal, it is always best to maximize: generalizability with respect to populations (A), precision in control and measurement of variables related to the behavior of interest (B) and existential realism for the participants and of the context in which behavior is observed (C). For example, if one uses a controlled laboratory experiment to maximize precision in control and measurement of variables (B), one risks low generalizability to other populations and lack of realism with respect to the work context. It is simply not possible to use a single study that addresses all the possible threats to validity.

The solution is triangulation: using more than one research approach to address the same question. Triangulation sometimes refers to the use of different methods within one paradigm, such as using different methods of operationalizing behavior for different experiments. We argue that triangulation across scientific and design disciplines is more likely to be beneficial, especially in interdisciplinary fields such as HCI.

Unfortunately, individual researchers cannot be expert in all the component disciplines and rarely have the resources to conduct a wide range of different studies. On the other hand, research laboratories such as CENA that investigate complex real-world problems generally employ researchers and designers with different backgrounds. Researchers in such environments can, by relating their work to each other, triangulate across disciplines and avoid many of the problems inherent in single-discipline based research. Addressing individual problems with multiple methods should produce results that are significantly more robust and useful.

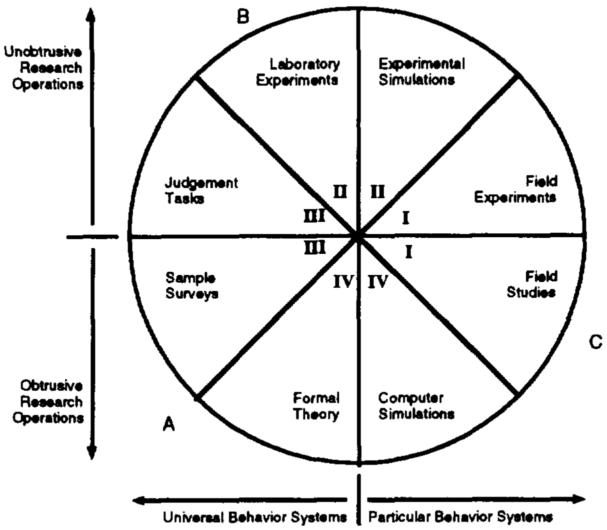

Figure 12: The eight major research strategies can be related to one another along several dimensions. Source: Runkel and McGrath (1972).

I.	Settings in Natural systems
II.	Contrived and created settings
III.	Behavior not setting dependent
IV.	No observation of behavior required

A.	Point of maximum concern with generality over actors.
B.	Point of maximum concern with precision of measurement of behavior.
C.	Point of maximum concern with system character of context.

The framework presented in Figure 5 attempts to relate the wide variety of research activities in HCI, both scientific and design. If we can avoid creating a "Tower of Babel" and learn to talk to each other, we can benefit from each other's work and find more productive solutions to our research problems.

CONCLUSIONS

Designers are faced with a bewildering array of different methods and philosophical perspectives. Bellotti et al. (1995) address the question of whether or not the science base of HCI has anything to contribute to real design. They conclude that theoretically-grounded HCI techniques can be effective, but only if the "end-user requirements of the design practitioners are properly understood, and the value of such techniques can be demonstrated". Showing how these techniques, along with their underlying assumptions, relate to standard design practices, should help to achieve this goal.

Individual research strategies are necessarily limited and contain serious flaws. Yet what is a weakness for individual research paradigms is actually a strength for HCI: if we can understand the relationships among them, we can triangulate across our component disciplines and improve the validity and value of our research.

The framework we present here seeks to provide a simple demonstration of how scientific and design disciplines in HCI relate to each other. By presenting the component disciplines in the context of interaction between users and computers, we can re-examine our own research setting and clarify how the different kinds of work can benefit from each other. The illustration of the framework with projects from a real-world research setting, CENA, shows how, despite its simplicity, the framework covers a wide range of research and design activities. Our goal is not only to describe, but also to suggest fruitful directions for additional research. By triangulating across the scientific and design disciplines that compose HCI, we can increase the validity of our research results and, we hope, make them more useful to real-world problems.

ACKNOWLEDGEMENTS

Merci à l'Équipe 9W! Lotte Bailyn provided the original insight about the relationship between the theoretical and empirical levels. Also, our thanks to Michel Beaudouin-Lafon, Marie Christine Bresolle, Dominique Colin De Verdière, Stéphane Chatty and Christophe Mertz for their comments on earlier drafts and to Marcel Leroux for his support.

REFERENCES

Alexander, C., Ishikawa, S., and Silverstein, M. (1977) *A Pattern Language.* New York: Oxford University Press.

Barley, S. R., Meyer, G.W., Gash, D.C. (1988). Cultures of Culture: Academics, Practitioners and the Pragmatics of Normative Control. *Administrative Science Quarterly,* Vol. 33, pp. 24-60.

Bellotti, V., Buckingham Shum, S., MacLean, A., Hammond, N. (1995) Multidisciplinary modelling in HCI design...in theory and in practice. In *Proceedings of CHI '95 ACM Conference on Human Factors in Computing* (pp. 429-436) Boulder, Colorado: ACM Press.

Bentley, R., Hughes, J.A., Randall, D., Rodden, T., Sawyer, P., Shapiro, D. and Somerville, I. (1992) Ethnographically-informed systems design for air traffic control. In *Proceedings of CSCW '92, ACM Conference on Computer-Supported Cooperative Work.* (pp. 122-129) Toronto, Ontario: ACM Press.

Bressolle, M.C., Pavard, B., and Leroux, M. (1995) The role of multimodal communicaton in cooperaton and intention recognition: The case of Air Traffic Control. In *CMC'95, The International Conference on Cooperative and Multimodal Communication: Theory and Applications.* Eindhoven, The Netherlands.

Card, S., Moran, T., and Newell, A. (1983) *The Psychology of Human-Computer Interaction.* Hillsdale, N.J.: Lawrence Ehrlbaum Associates.

Carroll, J. (1995) Scenario-based design. Envisioning work and technology in system development. NY: Wiley & Sons.

Chalmers, A.F. (1976) *What is this thing called Science? An Assessment of the Nature and Status of Science and its methods.* St.Lucie: University of Queensland Press.

Chatty, S. & Lecoanet, P. (1996) Pen Computing and Air Traffic Control. *In Proceedings of CHI'96 ACM Conference on Human Factors in Computing,* (pp. 87-94) Vancouver, British Columbia: ACM Press.

Cook, T. and Campbell, D. (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Boston, MA: Houghton-Mifflin Company.

DNA (January, 1997) *La Lettre de la Direction de la Navigation Aérienne.* No. 22.

Fleming, M. and Levie, W.H. (1978) *Instructional Message Design.* Englewood Cliffs, NJ: Educational Technology Publications.

Gillies, D. (1993) *Philosophy of Science in the 20th Century.* Oxford: Blackwell Press.

Gras, A., Moricot, Poirot-Delpech, S.L., and Scardigli, V. (1994) *Faced with automation: The pilot, the controller and the engineer.* Publications of the Sorbonne: Paris.

Harper, R., Hughes, J., and Shapiro, D. (1991) Harmonious working and CSCW: Computer Technology and Air Traffic Control. In *Studies in CSCW: Theory, Practice and Design.* Bowers, J. and Bedford, S., Eds. North Holland: Amsterdam. pp. 225-235.

Hartwig, F. and Dearing, B.E. (1979) *Exploratory Data Analysis.* Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-001. Beverly Hills, CA: Sage Publications.

Heath, C. and Luff, P. (1991) Collaborative Activity and Technological Design: Task Coordination in the London Underground Control Rooms. In *Proceedings of ECSCW'91, The European Conference on Computer-Supported Cooperative Work.* Kluwer Press.

Hopkin, V.D. (1995) *Human Factors in Air Traffic Control.* London: Taylor & Francis.

Hughes, J.A., Randall, D. and Shapiro, D. (October, 1992) Faltering from Ethnography to Design. In*Proceedings of CSCW '92, ACM Conference on Computer-Supported Cooperative Work.* (pp. 115-122) Toronto, Ontario: ACM Press.

John, B. and Packer, H. (1995) Learning and using the Cognitive Walkthrough method: A case study approach. In *Proceedings of CHI '95 ACM Conference on Human Factors in Computing* (pp. 429-436) Boulder, CO: ACM Press.

Kuhn,T.S. (1962)*The Structure of Scientific Revolutions.* Chicago, IL: University of Chicago Press.

Kyng, M. & Greenbaum, J. (1991) *Design at Work.* Hillsdale, NJ.: Lawrence Ehrlbaum.

Latour, B. (1987). *Science in Action.* Cambridge, MA: Harvard University Press.

Leroux, M. (1993a) The role of expert systems in future cooperative tools for air traffic controllers. In *Proceedings of the 7th International Symposium on Aviation Psychology.* (pp. 26-29). Columbus, OH.

Leroux, M. (1993b) ERATO: Vers une nouvelle philosophie de coopération homme-machine dans l'ATC. In *Le Transpondeur,* No. 9.

Lo See, J.A., (1993) *A Historical Introduction to the Philosophy of Science.* Oxford: Oxford University Press.

Mackay, W. (1990) *Users and Customizable software: A Co-Adaptive Phenomenon.* Doctoral Dissertation, The Massachusetts Institute of Technology. Cambridge, MA.

Mackay, W. (March 1995). Réalité Augmentée : le meilleur des deux mondes. *La Recherche,* numéro spécial (No. 285) *L'ordinateur au doigt et à l'œil.*

Mackay, W. and Bødker, S. (1994) Workshop on Scenario-Based Design. In *CHI'94 Conference Companion.,* Boston, MA: ACM Press.

MacLean, A., Young, R., Bellotti, V. and Moran, T. (1991) Questions, Options and Criteria: Elements of design space analysis. *Human-Computer Interaction.* Vol. 6, No. 3&4. pp. 201-250.

McGrath, J., Martin, J. & Kulka, J. (1982) *Judgement Calls in Research.* CA: Sage Publications.

Mayhew, D.J. (1992) *Principles and Guidelines in Software User Interface Design.* Englewood Cliffs, NJ: Prentice Hall.

Mertz, C. (1996) *Résultats de l'expérimentation CINDI. CENA* Technical Report, NR96-304.

Nielsen, J. and Phillips, V.L. (1993) Estimating the relative usability of two interfaces: Heuristic, format and empirical methods. In *Proceedings of INTERCHI '93 ACM Conference on Human Factors in Computing* (pp. 214-221) Amsterdam, The Netherlands: ACM Press.

Norman, D.A. and Draper, S.W. (1986). *User-Centered System Design.* Hillsdale, New Jersey: Ehrlbaum Associates.

O'Hew, A. (1989) *An Introduction to the Philosophy of Science.* Oxford: Clarendon Press.

Popper, K.R. (1963) *Conjectures and Refutations.* Routledge & Kegan Paul.

Preux, F. (1994) Rôle des strips dans l'activité des contrôleurs. *Sélection Professionnelle IEEAC.* CENA.

Runkel, P. and McGrath, J. (1972) *Research on Human Behavior: A Systematic Guide to Method.* NY: Holt, Rinehart & Winston.

Smith, S., and Mosier, J. (1986) *Guidelines for Desiging User Interface Software.* Report ESD-TR-86-278, The MITRE Corporation, Bedford, MA.

Sidman, M. (1960). *Tactics of Scientific Resarch: Evaluating Experimental Data in Psychology.* New York, NY: Basic Books.

Simon, H.A. (1969) *The Sciences of the Artificial.* Cambridge, MA: MIT Press.

Suchman, L. (1987). *Plans and Situated Actions.* Cambridge, England: Cambridge University Press.

Wellner, P., Mackay, W. and Gold, R. (1993) Computer-Augmented Environments: Back to the Real World. Special issue of *Communications of the ACM,* Vol. 36, No.7.

Zorola-Villarreal, R., Pavard, B., and Bastide, R. (1995) SIM-COOP: A Tool to analyse and predict cooperation in complex environments. A Case Study: The introduction of a datalink between controllers and pilots. In the *Proceedings of the 5th International Conference on Human-Machine Interaction and Artificial Intelligence in Aerospace,* IHM-AI-AS 95, Toulouse, France.

Zuboff, S. (1988). *In the Age of the Smart Machine.* New York: Basic Books.