CHAPTER 5

Modern Theories

In the late 1980s, it became increasingly apparent that the early attempts at importing cognitive psychology theories were failing to live up to their expectations. Several researchers began to reflect on why this was the case, especially why they were unable to be more widely applied to the problems of design and computer use (e.g., Long and Dowell, 1996). There was much navel-gazing, cumulating in the realization that classical cognitive theories were inadequately formulated for informing system design (see Carroll, 1991). A number of problems were identified, including that the theories were too low level, restricted in their scope and failed to deal with real world contexts (Barnard, 1991). The failings of a "one-stream" approach, whereby it was assumed that mainstream theory provided by pure science (i.e., cognitive psychology) could trickle down into the applied science of designing computer systems were exposed (see Long and Dowell, 1996). There was even criticism that psychologists were merely using the field of HCI as a test bed for trying out their general cognitive theories (Bannon and Bødker, 1991) or for validating the assumptions behind specific models (Barnard and May, 1999).

It was still hoped that theory could make a valuable contribution in HCI. The question raised by all this introspection, however, was what kind of theory and what role should it play? Several prominent researchers began to push for other approaches. Long and Dowell (1989, 1996), for example, made persistent calls for more domain-specific theories that could focus on the concerns of users interacting with computers to enable them to work effectively. Carroll et al. (1991) argued that users and designers would benefit more if the process by which tasks and artifacts co-evolved could be "better understood, articulated and critiqued" (p99).

Others revised and adapted their cognitive frameworks to be more representative and build directly on the concerns of HCI (e.g., Draper, 1992). New cognitive theories emerged that focused on interactivity rather than solely modeling what was assumed to happen "inside the head." It was recognized that a more appropriate conceptualization of cognition for HCI was one that was distributed across people, technologies and the environment and externalized. A central focus was the interplay between external representations and internal representations at the interface (e.g., Green et al., 1996; Hutchins, 1995; Kirsh, 1997; Scaife and Rogers, 1996; Wright et al., 2000).

There were also attempts to look for different theories that took into account how the environment affected human action and perception. Several ideas from ecological psychology were imported into HCI (e.g., Gaver, 1991; Norman, 1988). Other researchers began looking elsewhere for theories that were more encompassing, and which could address the concerns of interacting with computers in real-world contexts. By changing the boundaries of what was studied, and by looking at the phenomena of interest with different theoretical lenses and methods, it was assumed that a new set of research questions could be framed, which, in turn, could feed into the design of more usable computer artifacts (Bannon and Bødker, 1991).

Concomitantly, there was a "turn to the social" (Button, 1993): sociologists, anthropologists and others in the social sciences joined HCI, bringing with them new frameworks, theories and ideas about technology use and system design. Human-computer interactions were conceptualized as social phenomena (e.g., Heath and Luff, 1991). Most notable was the situated action (SA) approach and ethnography. A main thrust of the SA approach was to examine the *context* in which users interact with technologies: or put in social terms, how people use their particular circumstances to achieve intelligent action.

The approach known as ethnomethodology (Garfinkel, 1967; Garfinkel and Sacks, 1970) provided much of the conceptual and methodological underpinning for the early ethnography in HCI (Button, 1993). It offered new ways of describing the informal aspects of work, i.e., "the hurly burly of social relations in the workplace and locally specific skills required to perform any task" (Anderson, 1994, p154).

Table 5.1 shows each of the Modern theoretical approaches that are covered, outlining their origins, appropriation in HCI, their impact on research and practice and an "in a nutshell" description. Their selection and amount of coverage is based on their influence in HCI. A section on CSCW theories is also included to show how theory was instrumental in the establishment of that field.

Table 5.1: Modern Theoretical Approaches in HCI

Alternative cognitive approaches
5.1 External cognition
5.2 Distributed cognition
5.3 Ecological psychology
Social approaches
5.4 Situated action
5.5 Ethnomethodology and ethnography
5.6 CSCW theories
Other imported approaches
5.7 Activity theory
5.8 Grounded theory
5.9 Hybrid theories

5.1 EXTERNAL COGNITION

Larkin and Simon's (1987) classic cognitive science paper on why a diagram may be worth a thousand words became a landmark in HCI because it offered the first alternative computational account of cognition which focused on how people interact with external representations. Their seminal idea was that cognition be viewed as the interplay between internal and external representations, rather than only be about modeling what was assumed to happen inside the head. It was regarded by those who had become disaffected by cognitive models as a source of inspiration for rethinking HCI and for me, personally, provided an "aha" moment, leading to the development of a new theory of external cognition (see Rogers, 2008a).

Larkin and Simon's theoretical account made an important distinction between two kinds of external representation: diagrammatic and sentential representations. While being *informationally* equivalent they were considered to be *computationally* different. That is they contain the same information about the problem but the amount of cognitive effort required to come to the solution differed. They proposed that solutions to problems could be "read off" from diagrams that were implicit in sentences. People can readily switch their attention from one component to another in a diagram to draw conclusions in ways that are impossible to do with a sequence of sentences. Diagrams provide simultaneous information about the location of components in a form that enables objects and their relations to be easily tracked and maintained. From this, we can deduce that the best diagrams are those that make it obvious where to look to draw conclusions.

Larkin and Simon's paper paved the way for HCI researchers to begin in earnest to theorize the role of external representations in human-computer interactions. There was a palpable buzz in the early 1990s as they endeavored to change the face of theorizing in HCI. O'Malley and Draper (1992) proposed a display-based account that differentiated between the knowledge users need to internalize when learning to use display-based word processors (e.g., Word) and the knowledge that they can always depend upon being available in the external display. Norman (1993) had a big impact, popularizing the notion that knowledge resides both in "the head" and in "the world." Wright et al. (2000) developed a resource model that analyzed internal (e.g., memorized procedure) and external representations (e.g., written instructions). Kirsh (1997) developed a theory of interactivity that stressed how cognition can be extended in a variety of ways in what we can do, and allowing us to think more powerfully As well as reducing the cognitive effort that is needed to perform tasks, he argued that we should reframe external representations in terms of how they can enhance cognitive power. He suggested a number of ways, including providing a structure that can serve as a shareable object of thought; creating persistent referents; facilitating re-representation and the computation of more explicit encoding of information and helping to coordinate thought. A core aspect of interactivity is the ability to project structure onto things and then modify the world to materialize or reify that projection. People often reorder or rearrange objects in the environment, such as shuffling the letters around in a Scrabble tray to help them work out the best word given their set of letters (Maglio et al., 1999). Kirsh (2010) also stresses how we are always creating external representations; on the one hand they can help reduce memory load and the cognitive cost of computational tasks but, equally, they can do and allow us to think more powerfully.

The theory of external cognition was developed to systematically inform how new technologies, such as animations, multi-media and virtual reality could extend and enhance cognition

(Rogers and Scaife, 1998; Scaife and Rogers, 1996). A number of core dimensions were identified that could be used to guide the design of different kinds of external representations that would be of "added" cognitive value for particular users, domains and tasks. It suggested how interactive mechanisms enabled by computer technologies could be exploited to guide and scaffold learners in knowing where to look in order to interpret, make inferences and connections between the different elements of a graphical representation.

External Cognition in a Nutshell

A central property of external cognition is *computational offloading* — the extent to which different external representations vary the amount of cognitive effort required to carry out different activities (Scaife and Rogers, 1996). This is broken down into specific design dimensions, intended to guide the design of interactive representations. They include *re-representation* (how different external representations, that have the same abstract structure, make problem-solving easier or more difficult) and *graphical constraining* (how elements in a graphical representation are able to constrain the kinds of inferences that can be made about the underlying represented concept). The dimensions were further characterized in terms of design concepts with the purpose of framing questions, issues and trade-offs. Examples of design concepts are *cognitive tracing*, which refers to the way users are allowed to develop their own understanding and external memory of a representation of a topic by being allowed to modify and annotate it; *explicitness* and *visibility* which refers to how to make more salient certain aspects of a display such that they can be perceived and comprehended appropriately. Another design concept is *dynalinking*, which refers to how abstract representations, such as diagrams, are linked together with a more concrete illustration of what they stand for, such as a simulation. Changes in one are matched by changes in the other, enabling a better understanding of what the abstraction means.

The set of external cognition concepts were intended to suggest to designers ways of generating possible functions at the interface. For example, Masterman and Rogers (2002) developed a number of online activities that allowed children to create their own cognitive traces when learning about chronology using an interactive multimedia application. They have also been used for deciding how to design and combine interactive external representations for representing difficult subjects, such as dynamical systems in biology, chronology in history, the working of the cardiac system and crystallography (e.g., Gabrielli et al., 2000; Masterman and Rogers, 2002; Otero, 2003; Price, 2002). Sutcliffe (2000) has also shown how he used the theory to inform the design of multimedia explanations. The approach was also applied in work settings, to inform the design of online graphical representations that could facilitate and support complex distributed problem solving (Rodden et al., 2003; Scaife et al., 2002). Dynalinking has been used in a number of areas to explicitly show relationships among multiple dimensions where the information to be understood or learned is complex (Sutcliffe, 2000). For example, it has been used to represent complex data using various interactive visualizations, for domains like learning science subjects, economic forecasting, molecular modeling, and statistical analyses.

Other analytic frameworks that were developed under the umbrella of the external cognition approach include Green's (1989) cognitive dimensions and Wright et al.'s (2000) resource model.

5.1.1 COGNITIVE DIMENSIONS

Cognitive dimensions were intended to enable psychologists and importantly, designers, to make sense of and use when talking together about design issues. Green's overarching goal was to develop a set of high-level concepts that were both valuable and easy to use for evaluating the designs and assessment of informational artifacts, such as software applications. An example dimension is "viscosity," which simply refers to resistance to local change. The analogy of stirring a spoon in treacle (high viscosity) versus milk (low viscosity) quickly gives the idea. Having understood the concept in a familiar context, Green then showed how the dimension could be further explored to describe the various aspects of interacting with the information structure of a software application. In a nutshell, the concept is used to examine "how much work you have to do if you change your mind" (Green, 1990, p79). Different kinds of viscosity were described, such as "knock-on' viscosity," where performing one goal-related action makes necessary the performance of a whole train of extraneous actions. The reason for this is due to constraint density: the new structure that results from performing the first action violates some constraint, which must be rectified by the second action, which in turn leads to a different violation, and so on. An example is editing a document using a word processor without widow control. The action of inserting a sentence at the beginning of the document can have a knock-on effect whereby the user must then go through the rest of the document to check that all the headers and bodies of text still lie on the same page.

The approach was meant to be broad-brush, and importantly, comprehensible to and usable by non-specialists. The original set of terms comprised a small vocabulary of about 12 terms that describe aspects of user interaction that are cognitively relevant. Besides viscosity, they included premature commitment ("are there strong constraints in terms of the order of how tasks are to be carried out?"), diffuseness ("how much space does the notation require to produce a certain result or express a meaning?") and visibility ("how readily can required parts of the notation be identified, accessed and made visible?"). Some are more intuitive to understand than others. One of Green's claims about the value of cognitive dimensions is that by identifying different kinds of dimensions at a suitable level of abstraction across applications, solutions found in one domain may be applicable to similar problems found in others.

Although never widely used, the lingua franca of "cog dims" has been influential. In particular, it has been used to determine why some interfaces are more effective than others. These include educational multimedia (e.g., Oliver, 1997; Price, 2002), collaborative writing (Woods, 1995), tangible user interfaces (Edge and Blackwell, 2006) and programming environments (Modugno et al., 1994; Yang et al., 1995). Kadoda et al. (1999) and Blackwell and Green (2000) also extended the approach by developing a generalized questionnaire in which the definitions of cog dims are provided for users rather than designers, who decide for themselves the features of a system that they wish to criticize. Designers and researchers who have been exposed to them for the first time have found them comprehensible, requiring not too much effort to understand and to learn how to use (Green et al., 1996). Indeed, when one first encounters the cog dims there is a certain quality about them that lends to articulation. They invite one to consider explicitly trade-offs in design solutions that might otherwise go unnoticed and which, importantly, can be traced to the cognitive phenomena they are derived from.

5.1.2 WRIGHT ET AL.'S RESOURCES MODEL

Wright et al. (2000) modeled external cognition in terms of *resources* that are drawn upon during user interaction. They categorized these in terms of plans, goals, possibilities, history, actions-effect relations or states. They could be represented internally (e.g., memorized procedure) or externally (e.g., written instructions). Configurations of these resources, distributed across internal and external representations, were assumed to be what informs an action. In addition, the way the resources are configured in the first place, is assumed to come about through various "interaction strategies." These include things like plan following and goal matching. Thus, a user's selection of a given action may arise through an internal goal matching strategy (e.g., delete the file) being activated in conjunction with an external "cause-effect relation" being perceived, (e.g., a dialog box popping up on the screen saying 'are you sure you want to delete this file?').

Wright et al.'s (2000) analytic framework identified patterns of interaction together with the variability of resources that are used at different stages of a task — such as determining when a user can depend on the external resources (e.g., action-effect relations) to constrain what to do next and when they must rely more on their own internal resources (e.g., plans, goals and history of actions). The idea was that the analyst could reflect on the problems with a given interface in terms of the demands the various patterns of resources place on the user.

What impact has external cognition had on HCI?

One of the main uses of the external cognition approach in HCI has been to enable researchers and designers to articulate designs and phenomena in terms of a set of core properties and design dimensions — which they did not have access to before. In so doing, a language, couched in how people manipulate representations, interact with objects, etc., at an interface, was provided, helping researchers to select, articulate and validate particular forms of external representation in terms of how they could support various activities being designed for. Besides the originators of the theoretical frameworks, they have been used by a number of others to inform the design of various interfaces. Their emphasis on determining the optimal way of structuring and presenting interactive content with respect to the cognitive effort involved can be viewed as being generative Although largely superseded by contemporary theories that address a broader range of user aspects, the extended cognition approach still has much to offer in terms of helping designers select and create interactive visualizations, feedback and multi-modal representations.

5.2 DISTRIBUTED COGNITION

The distributed cognition approach considers cognitive phenomena in terms of individuals, artifacts, and internal and external representations (Hutchins, 1995). It provides a more extensive account compared with external cognition. Typically, it involves describing a "cognitive system," which entails interactions among people, the artifacts they use, and the environment they are working in. It was initially developed by Hutchins and his colleagues in the late 1980s and proposed as a radically new paradigm for rethinking all domains of cognition (Hutchins, 1995). It was argued that what was problematic with the classical cognitive science approach was not its conceptual framework *per se,* but its exclusive focus on modeling the cognitive processes that occurred within one individual. Alternatively, Hutchins argued that what was needed was for the same conceptual framework to be applied to a range of cognitive systems, including socio-technical systems at large (i.e., groups of individual agents interacting with each other in a particular environment).

Part of the rationale for this extension was that, firstly, it was assumed to be easier and more accurate to determine the processes and properties of an "external" system — since they can arguably, to a large extent, be observed directly in ways not possible inside a person's head — and, secondly, they may actually be different and thus unable to be reduced to the cognitive properties of an individual. To reveal the properties and processes of a cognitive system requires doing an ethnographic field study of the setting and paying close attention to the activities of people and their interactions with material media (Hutchins, 1995). These are conceptualized in terms of "internal and external representational structures" (Hutchins, 1995, p135). It also involves examining how information is propagated through different media in the bounded cognitive system.

Distributed Cognition in a Nutshell

The distributed cognition approach provides an event-driven description of the information and its propagation through a cognitive system. The cognitive system might be one person's use of a computational tool, such as a calculator; two people's joint activities when designing the layout for the front page of a newspaper, using a shared authoring tool, or more widely, a large team of software developers and programmers, examining how they coordinate their work with one another, using a variety of mediating artifacts, such as schedules, clocks, to-do lists and shared files.

The granularity of analysis varies depending on the activities and cognitive system being observed and the research or design questions being asked. For example, if the goal is to examine how a team of pilots fly a plane — with a view to improving communication between them — then the focus will be on the interactions and communications that take place between them and their instruments, at a fine level of granularity. If the goal is to understand how pilots learn how to fly — with a view to developing new training materials — then the focus will be at a coarser grain of analysis, taking into account the cultural, historical, and learning aspects involved in becoming a pilot.

The description produced may cover a period of a day, an hour or only minutes, depending on the study's focus. For the longer periods, verbal descriptions are primarily used. For the shorter periods, micro-level analyses of the cognitive processes are meticulously plotted using diagrammatic forms and other graphical representations. The rationale for performing the finer levels of analysis is to reveal practices and discrepancies that would go unnoticed using coarser grains of analysis, but which reveal themselves as critical to the work activity.

A distributed cognition analysis typically involves examining:

- The distributed problem-solving that takes place (including the way people work together to solve a problem).
- The role of verbal and non-verbal behavior (including what is said, what is implied by glances, winks, etc. and what is not said).
- The various coordinating mechanisms that are used, e.g., rules, procedures.
- The various ways communication takes place as the collaborative activity progresses.

- How knowledge is shared and accessed.

It should be stressed that there isn't one single way of doing a distributed cognition analysis. Within work settings, data is collected and then analyzed and interpreted in terms of work practices, routines and procedures followed, and the work arounds that teams develop when coping with the various demands placed upon them at different times during their work. Breakdowns, incidents or unusual happenings are highlighted, especially where it is discovered that excessive time is being spent doing something, errors were made using a system, or a piece of information was passed on incorrectly to someone else or misheard.

Problems can also be described in terms of the communication pathways that are being hindered or the breakdowns arising due to information not propagating effectively from one representational state to another. This level of analysis can reveal where information is being distorted, resulting in poor communication or inefficiency. Conversely, it can show when different technologies and the representations displayed via them are effective at mediating certain work activities and how well they are coordinated.

Hutchins emphasizes that an important part of doing a distributed cognition analysis is to have a deep understanding of the work domain that is being studied. He recommends, where possible, that the investigators learn the trade under study. This can take a team of researchers several months and even years to accomplish and in most cases this is impractical for a research or design team to do. Alternatively, it is possible to spend a few weeks immersed in the culture and setting of a specific team to learn enough about the organization and its work practices to conduct a focused analysis of a particular cognitive system.

The distributed cognition approach has been used primarily by researchers to analyze a variety of cognitive systems, including airline cockpits (Hutchins and Klausen, 1996; Hutchins and Palen, 1997), air traffic control (Halverson, 1995), call centers (Ackerman and Halverson, 1998), software teams (Flor and Hutchins, 1992), control systems (Garbis and Waern, 1999), emergency rooms (Artman and Waern, 1999), emergency medical dispatch (Furniss and Blandford, 2006) and engineering practice (Rogers, 1993, 1994). One of the main outcomes of the distributed cognition approach is an explication of the complex interdependencies between people and artifacts in their work activities. An important part of the analysis is identifying the problems, breakdowns and the distributed problem-solving processes that emerge to deal with them. In so doing, it provides multi-level accounts, weaving together "the data, the actions, the interpretations (from the analyst), and the ethnographic grounding as they are needed" (Hutchins and Klausen, 1996, p19). For example, Hutchins' account of ship navigation provides several interdependent levels of explanation, including how navigation is performed by a team on the bridge of a ship; what and how navigational tools are used, how information about the position of the ship is propagated and transformed through the different media and the tools that are used.

As a theoretical approach, it has received considerable attention from researchers in the cognitive and social sciences, most being very favorable. However, there have been criticisms of the approach, mainly as a continuation of an ongoing objection to cognitive science as a valid field of study and, in particular, the very notion of cognition (e.g., Button, 1997). In terms of its application in HCI, Nardi (1996, 2002) has voiced her concerns about its utility in HCI. Her main criticism stems from the need to do extensive fieldwork before being able to come to any conclusions or design decisions for a given work setting. Furthermore, she points out that there is not a set of interlinked concepts that can be readily used to pull things out from the data. In this sense, Nardi has a point: the distributed cognition approach is difficult to apply, since there is not a set of explicit features to be looking for, nor is there a check-list or recipe that can be easily followed when doing the analysis. It requires a high level of skill to move between different levels of analysis, to be able to dovetail between the detail and the abstract. As such it can never be viewed as a "quick and dirty" prescriptive method. The emphasis on doing (and interpreting) ethnographic fieldwork to understand a domain means that at the very least, considerable time, effort and skill is required to carry out an analysis.

Where the distributed cognition framework can be usefully applied to design concerns, is in providing a detailed level of analysis which can provide several pointers as to how to change a design (especially forms of representation) to improve user performance, or, more generally, a work practice. For example, Halverson (2002) discusses how in carrying out a detailed level of analysis of the representational states and processes involved at a call center, she was, firstly, able to identify why there were problems of coordination and, secondly, determine how the media used could be altered to change the representational states to be more optimal. Hence, design solutions can start to emerge from a detailed level of analysis because the nature of the descriptions of the cognitive system is at the same level as the proposed design. In other words, the low-level nature of a distributed cognition analysis can be most useful at revealing the necessary information to know how to change a design, when it has been identified as being problematic.

There have also been various efforts to develop more applied distributed cognition methods that are more accessible and easier to apply. One in particular that has been used by a number of researchers is Distributed Cognition for Teamwork (DiCoT) — essentially a structured approach for analyzing work systems and teamwork (Blandford and Furniss, 2005; Furniss and Blandford, 2010). The approach draws on core ideas from DC theory and combines them with more practical aspects of contextual design (Beyer and Holtzblatt, 1998; Holtzblatt and Jones, 1993), that resulted in a comprehensive set of underlying themes and principles intended to guide researchers in knowing what to focus on when analyzing and interpreting data from workplace settings. Themes include physical layout, information flow, and the design and use of artifacts; principles include subtle bodily supports (for example, pointing on a screen while replying to someone who walks in and asks a question is part of the mechanism of remembering where they are in a task) and arrangement of equipment (e.g., where computers, printers, etc., are in an office determines who has access to and can interact with information). The themes and principles are intended to help researchers organize their field observations into a set of interdependent models that can help elicit insights about user behavior.

Contextual design

Contextual design (Beyer and Holtzblatt, 1998) is not a theory but an applied approach that was developed to deal with the collection and interpretation of ethnographic findings. It is only briefly mentioned here because it was an important component in the development of the applied DiCoT framework. It is concerned with explicating context and the social aspects of user-interaction and how to use this to inform the design of software. It focuses on how to progress layers of abstractions rather than bridging analysis and design through examining the detail of each. It is also much more prescriptive, promoting a process of transforming data into a set of abstractions and models. The outcome is a very hands-on method of applying research findings, that has proven to be highly successful, with many other practitioners having adopted and used it. Part of its attraction lies in its conceptual scaffolding; it offers a step-by-step approach with various forms to fill in and use to transform findings into more formal structures.

A benefit of bringing together the various strands of the DC literature is to provide a more structured framework that can help researchers and developers to identify the strengths and limitations of the current artifact designs. In so doing, it should enable them to reason systematically about how to re-design the work settings, in terms of considering new technologies, work practices, physical layout, etc. Others have also started to use it to analyze work practices, including software team interactions (Sharp and Robinson, 2008) and mobile healthcare settings (McKnight and Doherty, 2008).

What impact has the distributed cognition approach had on HCI?

The distributed cognition approach has been widely used in HCI to analyze existing practices and to inform new and redesigns by examining how the form and variety of media in which information is currently represented might be transformed and what might be the consequences of this for a work practice. Partially in response to the criticism leveled at the difficulty of applying the distributed cognition approach, Hutchins and his colleagues (Hollan et al., 2000) set an agenda for how it could be used more widely within the context of HCI. They proposed it was well suited both to understanding the complex networked world of information and computer-mediated interactions and for informing the design of digital work materials and collaborative work places. They suggested a comprehensive methodological framework for achieving this.

Conducting a detailed distributed cognition analysis and using the DiCoT method has enabled researchers and designers to explore the trade-offs and likely outcomes of potential solutions and in so doing suggest a set of requirements grounded in the details of the work place, e.g., types of information resources, that are considered suitable for specific kinds of activities. The way theory has been applied from the DC approach has been largely descriptive and, to a lesser extent, generative; providing a detailed articulation of a cognitive system, and in so doing, providing the basis from which to generate design solutions.

5.3 ECOLOGICAL PSYCHOLOGY

The ecological psychology approach — originally developed by Gibson (1966, 1979) — was also considered by several researchers to be more relevant to HCI than the classical cognitive theories, especially for addressing how users interacted with the external world. For Bill Gaver (2008), reading Gibson was a revelation; making a huge impact on his thinking and convincing him of the importance of *contextualizing* human computer interactions in the environment they occur in rather than following the mainstream cognitive approach of isolating and identifying representations solely in the head.

Gibson's view was that psychology should be the study of the interaction between humans and their environment. This involved describing in detail the environment and people's ordinary activities within it (Neisser, 1985). HCI researchers took his philosophy and insights to heart, adapting his concepts in order to examine how people interacted with technological artifacts (Gaver, 1991; Kirsh, 2001; Norman, 1988; Rasmussen and Rouse, 1981; Vicente, 1995; Woods, 1995).

Ecological Psychology in a Nutshell

A central part of Gibson's ecological psychology theory is the notion of invariant structures in the environment and how they relate to human perception and action. Two that were considered most relevant to HCI were *ecological constraints* and *affordances*. Ecological constraints refer to structures in the external world that guide people's actions rather than those that are determined by internal cognitive processes. An affordance refers to the relationship between the properties of a person and the perceptual properties of an object in the environment. Within the context of HCI, it is used to refer to attributes of objects that allow people to know how to use them. In a nutshell, to afford is taken to mean "to give a clue" (Norman, 1988). Specifically, when the affordances of an object are perceptually obvious it is assumed that they make it easy to know how to interact with the object (e.g., door handles afford pulling, cup handles afford grasping). Norman (1988) provided a range of examples of affordances associated with everyday objects such as doors and switches — that were easy to understand and use when talking about interfaces.

This explication of affordances in HCI is simpler than Gibson's original idea — and to some extent that has been part of its appeal. It rapidly came into widespread use in HCI providing a way for researchers and designers, alike, to describe interfaces, suggesting to the user what to do when carrying out a task. It provided them with an easy to use shared articulatory device, helping them think about how to represent objects at the interface that could readily afford permissible actions (Gaver, 1991) and providing cues as to how to interact with interface objects more easily and efficiently.

It is not necessary to know about the original ecological psychology theory to understand the concepts that have been imported into HCI. Instead, the idea of invariant structures is taken as a given. The problem of only having a shallow understanding of an affordance, however, is that it requires working out what are affordable objects at the interface (St. Amant, 1999). There are no abstractions, methods, rules or guidelines to help the researcher identify instances of something — only analogies drawn from the real world.

Indeed, many designers began to use the term affordances to apply to everything, and as a way of thinking and talking about what adding a feature to the interface might mean to the user. It became easy to slip into talking about the meaning of an icon, the way a scroll bar moved, and the positioning of a window — as being easy to understand, because they afforded clicking on. Norman, however, was horrified at how sloppily the term had become used in common design parlance. To better articulate how to use the notion of affordances at the interface, he thought it important to understand the distinction between two kinds: perceived and real (Norman, 1999). On the one hand, physical objects were considered to have real affordances, as described above, like grasping, which are perceptually obvious and do not have to be learned. User interfaces that are screen-based, on the other hand, do not have these kinds of real affordances. Importantly, this means that users have to learn the meaning and function of each object represented at the interface before knowing how to act. Norman argued that screen-based interfaces have perceived affordances, which are based on learned conventions and feedback. For example, having a red flashing button icon appear at the interface may provide visual cues to enable the user to perceive that clicking on that icon is a meaningful useful action at that given time in their interaction with the system, that has a known outcome.

Vicente (1995) and Vicente and Rasmussen's (1990) considered it more beneficial to import more of the original theory into HCI, and so developed the Ecological Interface Design framework. Affordances were described in terms of a number of actions (e.g., moving, cutting, throwing, carrying). The various actions are sorted into a hierarchy of categories, based on what, why and how they afford. The framework was intended to allow designers to analyze a system at different levels, which correspond to the levels in the hierarchy.

Kirsh (2001) also proposed operationalizing the notion of affordance by grounding it more in the original Gibsonian ideas. Instead of couching it in terms of objects giving clues as to what to do, he proposed viewing affordance in terms of structures in the environment that *invite* people to do something. The term he used was *entry points*. Consider the way information is laid out on posters, websites and magazines; they provide different entry points for scanning, reading and following. These include headlines, columns, pictures, cartoons, figures, tables and icons. Well-designed information allows a person's attention to move rapidly from entry point to entry point for different sections (e.g., menu options, lists, descriptions). Poorly designed information does not have clear entry points — it is hard to find things. In Kirsh's terms, entry points are affordances in the sense of inviting people to carry out an activity (e.g., read it, scan it, look at it, listen to it, click on it). This reconceptualization potentially has more design purchase as it encourages designers to think about the coordination and sequencing of actions and the kind of feedback to provide, in relation to how objects are positioned and structured at an interface — rather than simply whether an object *per se* suggests what to do with them.

The concept of entry points has since been used successfully in interaction design as a design and conceptual tool. For example, Lidwell et al. (2006) have operationalized it as a generative design principle, describing features that, on the one hand, lure people into them and, on the other, do not deter them from entering them. Rogers et al. (2009) used entry points as the basis of their Shared Information Spaces framework, which was intended for researchers to think about how to constrain or invite group participation and collaboration through the layout of a physical room, the display and device interfaces provided and the kind and way information is presented (physical or digital). The assumption is that the design of entry points can provide different ways for participants to collaborate in both verbal and physical modes that lend themselves to more or less equitable participation. Hornecker et al. (2007) have also incorporated entry points into their comprehensive framework on the shareability of devices. The framework also includes other concepts such as access points, overview and fluidity of sharing, which are intended to show the relationship between elements of space, technology and people, by denoting design characteristics that invite people into engagement with a group activity and entice them to interact and join a group's activity. The assumption is that considering the relationship between the various elements in the framework can encourage a more comparative approach to designing interfaces for shared use. In so doing, it can enable a number of more specific research questions and hypotheses to be generated that could be investigated experimentally.

What impact has ecological psychology had on HCI?

A main contribution of the ecological psychology approach in HCI has been to extend its discourse, primarily in terms of articulating certain properties about an interface or space in terms of their behavior, appearance and properties. As such the role of theory is largely descriptive, providing design concepts. Most significantly, it has generated core terms that have become part of interaction design's everyday parlance, namely affordance and entry points.

5.4 SITUATED ACTION

The turn "to the social" took place in the late 1980s and early 1990s as a reaction against the dominant cognitive paradigm in HCI. During that time, several sociologists proposed alternative approaches for analyzing user-interaction (see Button, 2003; Shapiro, 1994) that focused on the social aspects of work settings and technology support. Lucy Suchman's (1987) book *Plans and Situated Action* took the field by storm and was universally read by all in HCI; its impact was to have a profound effect on how computation, programming, users and interface design were construed and researched. As well as providing an indisputable critique of the classical cognitive approaches, her alternative ideas about situated action resonated with many who had become disaffected with information-processing models underlying much of HCI.

The situated action *approach* has its origins in cultural anthropology. Suchman (1987) argued for "accounts of relations among people, and between people and the historically and culturally constituted worlds that they inhabit" (p71). To achieve this requires examining the relationship between "structures of action and the resources and constraints afforded by physical and social circumstances" (p179). This involves analyzing "how people use their circumstances to achieve intelligent action (…) rather than attempting to abstract action away from its circumstances" (p50). Suchman was quite clear in her intentions not to produce formal models of knowledge and action, but to explore the relationship of knowledge and action to the specific circumstances in which knowing and acting happen.

Situated Action in a Nutshell

The situated action approach offers detailed accounts of how technology is used by people in different contexts, which can often be quite different from the way the technology was intended to be used. The method used to reveal these discrepancies is predominantly ethnographic (i.e., carrying out extensive observations, interviews, collecting video and note taking of a particular setting). Typically, the findings are contrasted with the prescribed way of doing things, i.e., how people ought to be using technology given the way it has been designed. Sometimes conversational analysis (CA) is used to interpret the dialogue and interactions that take place between users and machine. For example, one of the earliest studies, using this approach was Suchman's (1983) critique of office procedures in relation to the design of office technology. Her analysis showed how there is a big mismatch between how work is organized in the process of accomplishing it in a particular office and the idealized models of how people should follow procedures that underlie the design of office technology. Simply, people do not act or interact with technology in the way prescribed by these kinds of models. Instead, Suchman argued that designers would be much better positioned to design systems that could match the way people behave and use technology if they began by considering the actual details of a work practice. The benefits of doing so could then lead to the design of systems that are much more suited to the kinds of interpretative and problem-solving work that are central to office work.

In Suchman's (1987) much-cited study — of how pairs of users interacted with an expert help system, intended as a help facility for using with a photocopier — she stressed how the design of such systems would greatly benefit from analyses that focus on the unique details of the user's particular situation — rather than any preconceived models of how people ought (and will) follow instructions and procedures. Her detailed analysis of how the expert help system was unable to help users in many situations where they got stuck highlighted again the inadequacy of basing the design of an interactive system primarily on an abstract user model. In particular, her findings showed how novice users couldn't follow the procedures, as anticipated by the user model, but instead engaged in ongoing, situated interaction with the machine with respect to what they considered at that moment as an appropriate next action.

SA analyses have revealed that while people may have plans of action in mind, they often need to change them depending on what is actually happening in a specific situation. They use their embodied and past experiences to deal with the contingencies of the ongoing situation. The canonical example provided by Suchman (1987) is of someone going over the falls in a canoe.

"In planning to run a series of rapids in a canoe, one is very likely to sit for a while above the falls and plan one's decent. The plan might go something like "I'll get as far over to the left as possible, try to make it between those two large rocks, then back very hard to the right to make it around that next bunch." A great deal of deliberation, discussion, simulation, and reconstruction may go into such a plan. But however detailed, the plan stops short of the actual business of getting your canoe through the falls. When it really comes down to the details of responding to currents and handling a canoe, you effectively abandon the plan and fall back on whatever embodied skills are available to you."

Following Suchman, a number of field studies were published that explored the situated and social aspects of user interaction in the work contexts they occurred. The outcome was a corpus of detailed "thick" accounts of a diversity of work practices and based on these design guidance about the specifics of the setting studied (Plowman et al., 1995). However, a criticism leveled at the situated action approach is its focus on the "particulars" of a given setting, making it difficult to step back and generalize. For example, Nardi (1996) exclaims how in reading about the minutiae of a particular field study "one finds oneself in a claustrophobic thicket of descriptive detail, lacking concepts with which to compare and generalize" (p92). This suggests it can be difficult for those used to seeing the world in abstractions to conceptualise it at such a level of detail. In an attempt to overcome this limitation, Hughes et al. (1997) proposed a generalizable framework to help structure the presentation of ethnographic findings in a way that was intended to act as a bridge between fieldwork and "emerging design decisions." The abstractions are discussed in terms of three core dimensions intended to orient the designer to thinking about particular design problems and concerns in a focused way, that in turn can help them articulate why a solution might be particularly helpful or supportive. These are "Distributed Coordination" (work tasks are performed as patterns of activity, e.g., division of labor), "plans and procedures" (the organizational support for distributed coordination, such as project plans and schedules, job descriptions) and "awareness of work" (the organization of work activities that makes them "visible" to others doing the work).

What impact has the situated action approach had in HCI?

The influence of the situated action approach on HCI practice has been pervasive. Suchman became one of the most frequently cited authors in the HCI literature. It changed the way researchers thought of computer interactions and work activities, taking context to be a focal concern. Several researchers reported how the situated action approach has profoundly changed the way they think about how they conceptualise and develop system architectures and interface design (e.g., Button and Dourish, 1996; Clancey, 1993). The situated approach has also become part of designer's talk; concepts of "situatedness" and "context" often being mentioned as important to design for. Hence, the situated action approach has, arguably, had a considerable influence on designers. Nowadays, it is increasingly common for designers and others to spend time "in the field" understanding the context and situation they are designing for before proposing design solutions (Bly, 1997).

Its contribution is descriptive, providing accounts of working practices. It has also had a big impact in the field , facilitating the widespread use of socially oriented concepts, such as context, and inspiring the development of analytic frameworks.

5.5 ETHNOMETHODOLOGY AND ETHNOGRAPHY

Another significant contribution to the "turn to the social" was ethnomethodologically informed ethnography, where field studies were conducted of work practices and interpreted in terms of the practical accomplishment of the people involved (Anderson, 1994). Similar to the situated action approach, it was developed in HCI as a reaction against mainstream cognitive theories. As the name suggests, it is considered an approach to adopt within HCI rather than a theory *per se.*

Ethnomethodology was originally proposed as an alternative *methodology* in sociology, intended to replace traditional top-down theories that sought to identify invariant structures (Garfinkel, 1967; Garfinkel and Sacks, 1970). Such external points of view of the world were considered not at all representative of the actual state of affairs. In this sense, it has an anti-theoretical stance, being quite explicit about its epistemological origins.

Ethnomethodology in a Nutshell

Ethnomethodology is concerned with how people accomplish social order in their everyday and work settings. Social order refers to the interactional work through which people conduct themselves (Garfinkel, 2002). This is viewed as an accomplishment in how society's members craft their moment-to-moment interaction. Essentially, it views people as shaping their actions rather than their actions being shaped by their environment. Ethnographic data is collected and analyzed to reveal how this is achieved. The accounts of work practices are presented largely as thick descriptions (Geertz, 1993). By this it is meant extensive and very detailed accounts.

Within HCI, the ethnomethodological approach has provided illuminating accounts of the details of work practices through which actions and interactions are achieved. Hence, it is an approach rather than a theory. It was popularized mainly by British sociologists, who used it to analyze a number of workplace settings; the most well known were of a control center in the London Underground (Heath and Luff, 1991) and of air traffic control (Bentley et al., 1992). They can be very revealing, often exposing taken for granted working practices, which turn out to be central to the efficacy of how a technological system is being used in a setting.

They have also been used to evaluate a number of technology designs and interventions, including Heath and Luff's series of studies on media spaces, augmented paper and hospital equipment (cited in Button, 2003). A recent study of theirs has also used it, first, to examine the interactional and sequential organization of museum visitor's actions and second, based on the findings, to inform the design of the conduct of a robot guide (Yamazaki et al., 2009). The robot guide was provided with various resources to engage visitors when interacting with a particular art exhibit. An evaluation of the robot *in situ* revealed how these resources were useful in engaging visitors in explanations.

A tension ethnomethodologists have had to confront, when working in HCI, is to show how their accounts can be useful for the design of technology and work. To begin, there was an expectation that the rich descriptions would lend themselves to being translated into "design implications" and several, started to add them at the end of their descriptions of field studies. The problem of asking ethnomethodologists to venture into this unfamiliar territory — namely, offering advice for others to follow — however, is that it forces them to come up with a cursory set of bullet points (Rogers, 1997). Many felt uncomfortable offering advice to others, whose very profession is to design — which clearly theirs is not. Their role is regarded as descriptive not prescriptive (Cooper, 1991). By bowing to such pressure, the design guidance ended up being rather tokenistic and not in keeping with the rich descriptions. For example, in one study, Anderson et al. (1993) provided a very detailed and insightful descriptive account of an organization's working practice. Following this, they outlined four brief "bullet-point" guidelines. One of these is that designers need support tools that take up a minimal amount of their time and that such tools should be adaptive to the exigencies of changing priorities. Such an observation is stating the obvious and could have easily been recognized without the need of a detailed field study. It is not surprising, therefore, that this form of abstracting from detailed field studies opened itself up for criticism; "most designers know the former only too well and desire the latter only too much" (Rogers, 1997, p68).

Recognizing this as a dilemma resulted in rethinking what else could be offered besides bullet points tacked on the end of thick descriptions. As one alternative, Button and Dourish (1996) proposed a core set of social mechanisms that had been proposed by the founders of ethnomethodology. These included the higher order concepts of practical action, order, accountability and coordination. Furthermore, they argued that ethnomethodologists and interaction designers could benefit by trying to see the world through each other's perspective: "design should adopt the analytic mentality of ethnomethodology, and ethnomethodology should don the practical mantle of design" (p22). It was suggested that this form of synergism could be achieved through system design taking on board concepts, such as situatedness, practical action, order and accountability, while ethnomethodology could benefit from taking on board system design concepts like generalization, configuration, data and process and mutability. While this approach was a laudable attempt at pushing a new dialogue in HCI it never took off. This is partly because the concepts were too difficult to define, appropriate and use in the different contexts.

Besides ethnomethodological-ethnography, a number of other ethnographic field studies have been published in the HCI and CSCW literature. Their theoretical and analytic framings vary; some being anthropological ethnography, focusing on cultural aspects, and others being cognitive ethnography, using concepts from distributed cognition to frame research questions and explain their data. Such eclecticism is considered by most to be healthy for the field, providing multiple perspectives and ways of understanding a diversity of human computer interactions, besides only being viewed from an ethnomethodology stance. However, at a recent CHI conference, a controversial paper surfaced, arguing that there was only really one acceptable way of grounding ethnography in HCI, namely in ethnomethodology (Crabtree et al., 2009). Much criticism was pitted against the new ethnographies that have been published on everyday settings, and which are not concerned with systems design *per se*. Crabtree et al.'s beef with the new ethnographies is in the "methodological dangers" that might arise when switching the focus to cultural practice; the ensuing accounts become more of a *literary practice* (Crabtree et al.'s emphasis) where design ideas are based on "conceptual rhetoric" rather than on the "organized conduct of those who will ultimately use the technology." Compared with the rich descriptions that the ethnomethodological approach provides they argue that cultural ethnography has become a vehicle for"producing social and cultural texts." Instead of acting on behalf of the members, the researchers have become wordsmiths plying their theories and understandings of culture.

Simply put, what is behind the ethnomethodologist's fear is that the new ethnographies are not kosher. However, the uproar following the publication of Crabtree at al.'s "ethnography considered harmful" paper suggests otherwise. Just as cognitive ethnography has provided new insights and ways of thinking about work practice, user experience and technology design so, too, are culturally-inspired ethnographies providing fresh insights into the everyday practices and appropriation of a diversity of technologies. The debate over what is acceptable HCI practice has long passed and more constructive debate about what contributions the different kinds of ethnography have to offer is to be welcomed.

Theory-driven Ethnography in a Nutshell

As discussed above, Rogers (1997) noted how drawing implications for design both belittles the ethnographer's rich descriptive accounts while appearing tokenistic to the designers. Moreover, it is often ill-suited to the kinds of generalizations that are expected and readily made from the findings of other kinds of user studies and conceptual analyses in HCI. More recently, (Dourish, 2006, 2007) extended this argument, critiquing the "implications for design" further. He stresses how the expectation that has arisen in HCI — that ethnographic field studies need to enumerate specific "implications for design" in order to be relevant to HCI is misplaced. Instead of assuming a researcher can "go out and find facts lying around in the world, dust them off, and bring them home to inform, educate, and delight" he argues that ethnography is deeply relevant for design, but that its value is elsewhere. This can get lost, when only focusing on the value of the design implications presented (or not) at the end of a field study. "Such lists underplay the more radical implications that may be caught up in ethnographic work; indeed, if the ethnographer returns from the field with little more than the lesson that the object in question should be green, should fit in a handbag, and should run for at least three weeks on two AA batteries, then I might venture that there isn't much to the ethnography." (Dourish, 2007, p5).

Importantly, the *theoretical* work of ethnography needs to be brought out much more, given its interpretive, analytic practice (Dourish, 2007). Moreover, the practice of reading for theory needs to be more up front since it may prove to be where the really significant "implications" are. This can involve examining ethnographical materials of cultural practices and the anthropologist's theoretical interpretations of them, which may be produced outside of the domain of technology development. The detailed and rich accounts of human experience can be read outside the terms of the specific field study, and, arguably provide guidance that is generalizable to other contexts — something that the ethnomethodologists have struggled with because of their adherence to the specific details. It does, however, require that the HCI researcher immerse themselves in the thick of cultural anthropology, which for some may be a joy, while for others too much of a journey into the unknown. For those with the time and curiosity, it can open up new possibilities for thinking about the design space, user experience and technologies.

What impact have ethnomethodology and ethnography had on HCI?

Ethnomethodology and ethnographic approaches have had a big impact on HCI and CSCW, by providing many insightful detailed descriptions of work and everyday practices. While the locus of their value has been contested, in terms of whether they should be making a theoretical or empirical contribution, they continue to be considered a fundamental part of doing user studies. They have informed design in various ways; suggesting new prototypes, new work practices and the conduct of user's interactions with technologies, in terms of how best they can be accomplished, by following the way they are conducted in human-human interactions and actions. Less successful, have been attempts to enrich design discourse, through providing a new lingua franca of concepts.

5.6 CSCW THEORIES

The field of CSCW emerged in the late 1980s, when a critical mass of researchers proposed an agenda for the social and organizational aspects of computing. The framing was concerned primarily with how computer technologies could be designed to support collaborative working practices. A particular interest was on how computer support could be developed to ameliorate the negative aspects of group working while enhancing or extending the positive aspects.

Unlike HCI, where cognitive psychology was largely the dominant theory that helped shape the field, initially a number of disparate theories were drawn upon from the social sciences, including organization, sociological and social psychology. They were primarily concerned with how organizations, groups and teams work together, but were approached from different levels and granularity of analysis, methodological stances and epistemologies. For example, theories arising from sociology included Gerson and Star's (1986) articulation work together with approaches stemming from Marxism, such as the division of labor, that adopted a critical stance towards how work is accomplished; theories coming from organizational psychology, included sociotechnical systems, with a focus on job satisfaction and work design principles, such as Enid Mumford's classic ETHICs (effective technical and human implementation of computer systems) method (see Mumford and Weir, 1979; and from social psychology, a number of group theories about social phenomena were cherry picked, such as why groups exert less effort than individuals (see McGrath, 1984).

A perusal of the early CSCW literature suggests that many more theories were brought in from a variety of perspectives, as CSCW established itself, compared with how HCI evolved. Part of the reason for this is that the study of the social covers a much broader range of interactions, concerns and topics — from why people behave differently in groups than by themselves to how organizations manage their businesses, and hence there were more disciplines and theories to draw upon by the diversity of researchers attracted to the emerging field. Arguably, this denser level of theorizing resulted in a wider range of applications and systems being analyzed, informed and generated (e.g., decision-making systems, workflow systems, computer-mediated communication, social networking).

It is simply not possible to do justice to all the theoretical developments in CSCW in one small section. Instead, I mention a few of them here to give a flavor of the different kinds that came to the fore and the ways they were applied. There are, of course, parallels with HCI theory development, in terms of how researchers saw their role in helping to account for, generate hypotheses, analyze and inform the design of user studies and computing systems. However, there are some differences, notably, in the scale and scope of the theories that address the complexity and uncertainty of how organizations and groups work, compared to the early attempts at predictability and controllability of individual cognition/user interaction.

Group theories were able to provide insights about social behavior and directions for how technologies might support their work activities more effectively. A classic was McGrath's (1991) theory and typology of group modes, known as TIP, that was influential in shifting from lab-based studies of groups to considering how they operate in real world settings (Grudin, 2008). He argued that groups engage in four essential modes of operation (inception, problem-solving, conflict resolution, and execution) for which there are three functions for each of these (production, group well-being and member support). Grudin (2008) notes how taking these all into account and their interdependencies results in a quite different way of viewing technology design and adoption; one that avoids the pitfalls of focusing exclusively on performance or productivity, which in, themselves, are likely to fail because of all the things that happen in groups, as identified by McGrath's typology.

Despite there being a large body of social psychology theories about small group behaviors, only a few were brought to the attention of researchers in CSCW (Kraut, 2003). The most well known was social loafing (individuals don't work as hard as when they work by themselves as they think the outcome of their efforts are being combined with those of others). Kraut (2003) has written extensively how these can be systematically applied to the design of group support systems, such as brainstorming tools and group decision-support tools. They have also been used to explain why groups behave differently or unexpectedly when participating in online groups and communities. An example he uses is the phenomenon of lurking or non-contribution in online communities, where people join an online community or other discussion group but only take on a passive role, reading what others had written and not contributing ideas or comments, themselves. Kraut (2003) suggests that as well as couching CSCW behaviors in these identified phenomena, *general* guidelines that have been derived from the theories in social psychology, such as social loafing (e.g., Karau and Williams, 1993), could be turned into *design* guidelines to suggest how to deal with them, by setting up technology in certain ways that would encourage or constrain more equal participation. In doing so, he shows how other factors and theories can be considered when trying to increase group participation levels to be more even, such as making the group appear attractive (using principles and theories of interpersonal attraction).

A controversial use of language theory

Theories of how people act through language, notably speech act theory, were used to develop the language/action framework, which in turn, was used to inform the design of a system, called the Coordinator, to help people work more effectively by improving the way they communicate with one another (Winograd and Flores, 1986). Speech act theory explains the functions utterances have in conversations (Austin, 1962; Searle, 1969). They can be direct (e.g., "I hereby declare you man and wife") or indirect (when someone says "It's hot in here" what they really mean is it OK to open the window). The Coordinator was designed to enable emails to be sent between people in the form of explicit speech acts. When sending someone a request, say "Could you get the report to me," the sender was also required to select the menu option "request." This was placed in the subject header of the message, thereby explicitly specifying the nature of the speech act. Thus, the Coordinator was designed to provide a straight forward conversational structure, allowing users to make clear the status of their work and, likewise, to be clear about the status of others' work in terms of various commitments.

The application of speech act theory in this manner, however, was subject to much criticism at the time by others in the research community who were incensed by the assumptions that speech act theory could be usefully applied to the design of a work system. Many heated debates ensued, often politically charged. A major concern was the extent to which the system *prescribed* how people should communicate. It was pointed out that asking users to specify *explicitly* the nature of their implicit speech acts was contrary to what they normally do in conversations. Moreover, forcing people to communicate in such an artificial way was regarded as highly undesirable. While some people may be very blatant about what they want doing, when they want it done by, and what they are prepared to do, most people tend to use more subtle and indirect forms of communication to advance their collaborations with others. The problem that Winograd and Flores came up against was people's resistance to radically change their way of communicating. Indeed, many of the people who tried using the Coordinator System in their work organizations either abandoned it or resorted to using only the free-form message facility, which had no explicit demands associated with it. It was asking too much of them to change the way they communicated and worked. However, it was successful in other kinds of organizations, namely those that are highly structured and need a highly structured system to support them. In particular, it has been much more successful in organizations, such as large manufacturing divisions of companies, where there is a great need for management of orders and where previous support has been mainly in the form of paper forms and inflexible task-specific data processing applications (Winograd, 1994).

Coordination theories were also influential, explaining the coordination work that is needed to enable groups to synchronize their efforts as a concerted action (e.g., Malone and Crowston, 1990). In an extensive review of his and others' work in this area, Schmidt (2011) expounded the coordination mechanisms that are integral to cooperative work, including the role played by constructs such as checklists, plans, blueprints, and operating procedures. Drawing from Marxism and other praxis-based theories, he established how material artifacts and practices are constructed, appropriated, applied and adapted, providing a theoretical basis from which to analyze coordination practices when new cooperative systems are being proposed or introduced into an organization.

Other theories that have been taken from critical sociology and the social sciences and shown how they can be applied in CSCW include actor network theory (see Latour, 2005; Law, 1987) and semiotic theory (De Souza, 2005). The former was developed as a reaction to the vague all encompassing terms prominent in sociology at the time, such as institutions, organizations, states and nations, replacing them with a more realistic and smaller set of associations for describing the very nature of societies, that encompasses both human individual actors and non-human, non individual entities. The latter draws on concepts from semiotics and computer science to investigate the relationship between designers and users, who are viewed as interlocutors in a communication process that takes place through the interface of words, graphics, and behavior.

5.7 ACTIVITY THEORY

Activity Theory (AT) is a product of Soviet Psychology that explained human behavior in terms of our practical activity with the world. It originated as part of the attempt to produce a Marxist Psychology, an enterprise usually associated with Vygotsky (e.g., 1962) and later Leontiev (e.g., 1978, 1989). In the last 30 years, versions of AT have become popular elsewhere, particularly in Scandinavia, Germany and now in the U.S. and UK. The newer "versions" of AT have been popular in research investigating "applied" problems, particularly those to do with work, technology and education.

Its conceptual framework was assumed to have much to offer to HCI, in terms of providing a means of analyzing actions and interactions with artifacts within a historical and cultural context (Bannon and Bødker, 1991; Bødker, 1989; Kuutti, 1996; Nardi, 1996). It first appeared in HCI in the late 1980s when Susanne Bødker (1989) applied it to the design of user interfaces for newspaper production. It was then brought to mainstream attention through her collaboration with Liam Bannon (Bannon and Bødker, 1991) where they showed how it could be used to analyze actions and interactions with artifacts within historical and cultural contexts. They argued that this kind of conceptual analysis could be used to inform the design of technologies that better suited workers in their work environments. Since their pioneering work, numerous edited volumes, case studies, PhD dissertations, and special journal issues have been published showing how AT can be adapted and applied to a diversity of areas, particularly the analysis of work, technology and education.

Besides Bødker and Bannon's seminal work, a number of researchers have promulgated its merits and value for HCI, notably, Yjro Engestrøm, Kari Kuutti, Olav Bertelsen, Wendy MacKay, David Redmiles and Jakob Bardram. But perhaps the most ardent and longstanding proponents are Bonnie Nardi and Victor Kaptelinin. Since the mid 1990s, they have tirelessly promoted the AT approach, arguing that it has much to offer HCI researchers and practitioners, especially compared with other cognitive and social approaches that have been imported into the field (Nardi and Kaptelinin, 2012). They claim it provides "the rigor and dedication of the scientific method of traditional cognitive science with the much needed attention to social and contextual factors necessary to HCI studies" (Kaptelinin and Nardi, 1997). Part of their mission has been to provide a broad framework for describing the structure, development and context of computer-supported activities that is easily usable by practitioners. This has included giving tutorials, workshops and an Activity checklist for identifying the most important factors influencing the use of computer technologies in a particular setting (Kaptelinin et al., 1999). Besides Nardi and Kaptelinin's reworking of AT for an HCI audience, several of the other AT researchers have elaborated and adapted Leontiev's (1978) original framework with applied goals in mind. Notable, is the highly cited work of Kuutti's (1996) extension of the hierarchical framework to show how information technology can be used to support different kinds of activities at different levels, and Nardi's (1996) adapted framework showing how it can be of value for examining data and eliciting new sets of design concerns. Nardi recast data from a field study that she had carried out earlier to compare the benefits of task-specific versus generic application software for making slides (Nardi and Johnson, 1994). In doing this exercise second time round, but with the added benefit of the conceptual framework of activity theory at hand, she found she was able to make more sense of her data. In particular, she cites how it enabled her to ask a more appropriate set of questions that allowed her subsequently to come up with an alternative set of recommendations about software architectures for the application of slide-making.

Activity Theory in a Nutshell

Activity Theory explains cultural practices (e.g., work, school) in the developmental, cultural and historical context in which they occur, by describing them in terms of "activities." The backbone of the theory is presented as a hierarchical model of activity that frames consciousness at different levels. These are operations, actions and activities. A number of principles are also proposed.

Focusing the analysis around the concept of an activity can help to identify tensions between the different elements of the system. An example of where it was used to show these was MacKay et al.'s (2000) study of users working with a new software tool that identified 19 shifts in attention between different parts of the tool interface and the task at hand. Some users spent so much time engaged in these shifts that they lost track of their original task. Using the theory helped the evaluators to focus on relevant incidents.

There are two key models: (i) an activity model and (ii) the mediating role of artifacts.

(i) The "classic" individual model (Figure 5.1)

At the bottom level of the model are operations, routinized behaviors that require little conscious attention, e.g., rapid typing. At an intermediate level are actions that are characterized by conscious planning, e.g., producing an index. The top level is the activity, and that provides a minimum meaningful context for understanding the individual actions, e.g., writing a chapter. There may be many different operations capable of fulfilling an action, and many actions capable of serving the same activity.

$$
\begin{array}{lcl}
\text{Activity} & - & \text{Motive} \\
\uparrow\downarrow & & \uparrow\downarrow \\
\text{Action} & - & \text{Goal} \\
\uparrow\downarrow & & \uparrow\downarrow \\
\text{Operation} & - & \text{Conditions}
\end{array}
$$

Figure 5.1: The original activity theory model.

Activities can be identified on the basis of the motives that elicit them, actions on the basis of conscious goals that guide them, and operations by the conditions necessary to attain the goals. However, there is an intimate and fluid link between levels. Actions can become operations as they become more automatic and operations can become actions when an operation encounters an obstacle, thus requiring conscious planning. Similarly, there is no strict demarcation between action and activity. If the motive changes then an activity can become an action. It is also important to realize that activities are not self-contained. Activities relate to others while actions may be part of different activities, and so on.

(ii) Mediating role of Artifacts

Artifacts can be physical, such as a book or a stone, or they can be abstract, such as a system of symbols or a set of rules. Physical artifacts have physical properties that cause humans to respond to them as direct objects to be acted upon. They also embody a set of social practices, their design reflecting a history of particular use. Leontiev (1981) describes the process of learning what these inherent properties are as one of appropriation, signifying the active nature of the learning that is needed. The kind of learning involved is one of identifying and participating in the activity appropriate to the artifact.

Consider an infant learning to feed with a spoon. Leontiev (1981) observes that, at first, the infant carries the spoon to its mouth as though it were handling any other object, not considering the need to hold it horizontal. Over time, with adult guidance, the spoon is shaped in the way it is because of the social practice — the activity — of feeding and, in turn, the infant's task is to learn that relationship — to discover what practice(s) the object embodies. By contrast a spoon dropped into the cage of a mouse, say, will only ever have the status of just another physical object — no different from that of a stone.

The idea of abstract artifacts follows from the idea of mediation, i.e., a fundamental characteristic of human development is the change from a direct mode of acting on the world to one that is mediated by something else. In AT, the artifacts involved in an activity mediate between the elements of it. The social context of an activity is also considered central. Even when seemingly working alone, an individual is still engaged in activities that are given meaning by a wider set of practices.

Engestrøm's (1990) extension of Activity Theory, known as "developmental work research" has also been influential in CSCW. His framework was designed to include other concepts (e.g., contradictions, community, rules and division of labor) that were pertinent to work contexts and which could provide conceptual leverage for exploring these. He widened the focus from the individual triangle of a single activity (subject, activity, and object) to include supra-individual concepts — tools, rules, community, and division of labor. By tools is meant the artifacts, signs, and means that mediate the subject and object; by community is meant those who share the same object; by rules is meant a set of agreed conventions and policies covering what it means to be a member of that community (set by laws, parents, managers, boards, etc.); and by division of labor is meant the primary means of classifying the labor in a workplace, e.g., manager, engineer, receptionist.

The extended versions allow consideration of networks of interrelated activities — forming an activity system. It has been used to analyze a range of work settings — usually where there is a problem with existing or newly implemented technology — providing both macro and micro level accounts. Several others have adopted Engestrøm's approach and have used the model to identify a range of problems and tensions in various settings. Some have taken this variant and adapted it further to suit their needs. These include Halloran et al.'s (2002) Activity Space framework for analyzing collaborative learning, Spasser's (2002) "realist" approach for analyzing the design and use of digital libraries and Collins et al.'s (2002) model employed to help identify user requirements for customer support engineers. One of the putative benefits from having a more extensive framework with a set of conceptual foci is how they structure and scaffold the researcher/designer in their analysis:

"We found that activity system tensions provide rich insights into system dynamics and opportunities for the evolution of the system." (Collins et al. op cit, p58).

The extended analytic frameworks have proven attractive because they offer a "rhetorical force of naming" (Halverson, 2002, p247), providing a set of terms that the analyst can use to match to instances in their data and, in so doing, systematically identify problems. However, it still relies largely on the analyst's interpretative skills and orientation as to what course to take through the data and how to relate this to which concepts of the framework. In some ways this is redolent of the problem discussed earlier concerning the application of cognitive modeling approaches to real world problems. There is little guidance (since it essentially is a subjective judgment) to determine the different kinds of activities — a lot depends on understanding the context in which they occur.

It is argued, therefore, that to achieve a level of competence in understanding and applying the various AT frameworks still requires considerable learning and experience (Rogers, 2008b). Hence, while, variants of the activity system model can be applied more readily, they are most useful for those who have developed them and understand activity theory in its historic context. When given to others not familiar with the original theory, their utility is arguably less and can even be problematic. For example, the basic abstractions of the model, like object and subject, were found to be difficult to follow, and easily confused with everyday uses of the terms when used by design and engineering teams (who were initially unfamiliar with them) to discuss user requirements (Collins et al., 2002).

AT does not provide a clear methodological prescription for the description or analysis of behavior as a set of procedures to be followed. Identifying elements in the framework is highly dependent on individual interpretation. One of the biggest problems with doing an AT analysis is working out when something should be described as a top-level activity and when something is better described as a lower-level action. For example, completing a software project is considered to be a top-level activity, while programming a module as an action. However, equally, programming a module could be viewed as an activity — if that was the object of the subject (person).

González (2006) tried to overcome this problem of distinguishing between levels by introducing a new intermediate concept to sit between an action and an activity and which describes "how tasks are aggregated and thematically connected on higher level units of work" (p53). He called this new level as one of engagements, which "thematically connect chains of actions towards the achievement of a purpose" (p9). Five types of engagements were outlined as specific units of work: requests, projects, problems, events and recurrents (p156). His idea behind analyzing actions/activities as types of work practices — rather than trying to decide whether to label them as actions or activities — is appealing since it can reveal more about what actually happens in the workplace. As part of the extended form of analysis, he suggested that the various actions that take place be viewed in relation to their higher-level purpose, such as a group manager composing an email and then sending it out to his team to motivate them. The emphasis is also on the way actions relate to other actions, rather than on how actions are performed through operations. The role of communication is also stressed in terms of how workers justify their motives and choice of which action/activity to follow at a given time.

Potentially, the outcome of performing this additional level of analysis is a richer interpretation of the field study data, and arguably a better understanding of how work gets accomplished on a moment-to-moment, what-to-do-next basis within the wider context of the purpose of the work. It switches the focus of the analysis from agonizing about the level at which to label something to examining the types of working spheres/engagements people have and pursue in terms of their temporal patterns, priorities and interdependencies with the work of others. It also enables a better linkage between the detailed ethnographic data collected in field studies and the conceptual labels of the framework.

What impact has Activity Theory had in HCI?

AT has been very popular, especially among Ph.D. students, as an explanatory framework. It has been used to couch and ground qualitative data in a variety of contexts. Numerous tensions and contradictions have been identified in workplace settings leading to the identification of specific needs for new technological tools. Its value has been in providing a structured framework that breaks down into a set of conceptual tools that can then be mapped onto features of complex, real-world contexts. In so doing, problems and opportunities for new interventions can be elicited. It has been popularized in Scandinavia, UK and the U.S.

5.8 GROUNDED THEORY

Grounded theory is not a theory *per se* but an approach that aims to help researchers develop theory from the systematic analysis and interpretation of empirical data, i.e., the theory derived is grounded in the data. Similar to AT, it has been a very popular choice amongst researchers wanting to make sense of the qualitative data they have collected, such as ethnographic video. The approach was originally developed by Glaser and Strauss (1967) and has been adopted and adapted by several researchers for different situations. Glaser and Strauss also individually (and with others) developed the theory in slightly different ways. Glaser (1992) documented the way the variants differ.

Grounded Theory in a Nutshell

The aim of grounded theory is to develop a theory that fits a set of collected data. In a nutshell, it is "a set of well-developed concepts related through statements of relationship, which together constitute an integrated framework that can be used to explain or predict phenomena" (Strauss and Corbin, 1998). To develop a "grounded" theory requires the researcher iteratively switching between data collection and data analysis. Initially, data is collected and analyzed to identify categories, then that analysis leads to the need for further data collection, which is analyzed, and more data is collected. Hence, data gathering is driven by the emerging theory and finishes when no further insights are gained from the alternating.

The goal of the grounded theory approach is to identify and define the properties and dimensions of relevant categories and then to use these as the basis for constructing a theory. There are essentially three kinds of coding:

(i) *Open coding* where categories, their properties, and dimensions are discovered in the data.

(ii) *Axial coding* where the categories are systematically fleshed out and related to their subcategories.

(iii) *Selective coding* where categories are refined and integrated to form a larger theoretical scheme.

Strauss and Corbin (1998) suggest collecting data that includes written records of analysis and diagrammatic representations of categories (which they call memos and diagrams). To help identify and characterize relevant categories, the researcher is encouraged to:

- question the data in order to generate ideas or consider different ways of looking at the data.
- analyze a word, phrase, or sentence, in order to understand better the meaning of an utterance, which in turn can trigger different ways of viewing the data.
- analyze the data through comparing objects or between abstract categories.

  Charmaz (2011) provides a practical guide in detail of the steps involved in coding.

In contrast to AT, there isn't a set of concepts or framework that can be used to make sense of the data. Instead, researchers need to draw on their own theoretical backgrounds. For example, when Sarker et al. (2001) used the grounded theory approach to develop a model of collaboration in virtual teams they drew from their background in the social sciences, using ideas from human conduct and social structure. Furniss et al. (2011) have used theoretical ideas from both distributed cognition and resilience engineering (a new way of thinking about safety, that enables organizations to create processes that are robust and flexible and which use resources proactively in the face of disruptions or ongoing production and economic pressures).

The types of questions researchers pose and iterate is key to which kinds of concepts they end up eliciting. Dourish et al. (2004a) used semi-structured interviews and grounded theory to examine how people answer the question "Is this system secure enough for what I want to do now?," in the context of ubiquitous and mobile technologies. This qualitative approach was used to explore the issues before moving on to develop more detailed questions, but their conclusions included suggested design modifications to take this perspective on security into account. Grinter (1998) carried out interviewing and observation at a number of sites as part of her qualitative approach to gathering data and developing a grounded theory about recomposition in software companies. She also engaged in participant observation, by helping to conduct usability studies, reviewing system architectures, facilitating project retrospectives, and process design.

Much depends on the skills of the analyst and their background, in knowing how to process the data using the layered coding scheme. It is also very time-consuming. Charmaz (2011) also emphasizes that, as the name suggests, developing theory is central to using Grounded Theory. Importantly, it provides a method for researchers to develop theory rather than apply existing theory. Its value for HCI lies in how a researcher can skillfully iterate between data collection and analysis to create a new theoretical understanding. This does not have to be a totally new theory replete with concepts and relations; but can comprise a frame work that demonstrates a hierarchy of classes and sub-classes for a given setting for which data has been collected.

What impact has the grounded theory approach had in HCI?

Grounded Theory has been used widely in HCI, providing insights into people's values, understanding and experience with technology (Furniss et al., 2011). It has become increasingly popular in interaction design to answer specific questions and design concerns (Grinter, 2011). It provides a method for generating new theory, by iterating between data collection and coding. This differs from other top-down approaches, which provide a theoretical framework to match instances or patterns found in the collected data against an existing set of concepts.

5.9 HYBRID THEORIES

To conclude this section, I briefly touch upon some of the attempts in HCI to synthesize a diversity of concepts from different theories and disciplines. The rationale for the over arching approaches was to provide more extensive and if possible, unified theories, compared with importing concepts arising from only one discipline. The late Leight Star (1996), for example, was very skilful at bringing together different strands of disparate theories to provide fresh insights. In one instance, she looked at similarities between Activity Theory and symbolic interactionalism (originating from American pragmatism) that showed links between them.

Pirolli and Card's (1997) information foraging food-theory (IFT) was also very insightful and provided a completely new way of thinking about searching on the web. In particular, it opened up the field of information visualization to many more researchers, leading to the development of new kinds of graphical representations and browsing tools. They describe searching for and making sense of information in terms of a number of concepts borrowed from evolution, biology and anthropology and classical information processing theory: "in many ways analogous to evolutionary ecological explanations of food-foraging strategies in anthropology and behavior ecology" (p5). The search strategies are viewed in terms of making correct decision points, which are influenced by the presence or absence of "scent." If the scent is strong enough, the person will make the correct choices; if not they will follow a more random walk.

The most overarching theories of HCI attempted to integrate theories from different fields at multiple levels of analysis. An ambitious example was Mantovani's (1996) eclectic model of HCI, that combined a diverse range of concepts and research findings from computer-supported cooperative work (CSCW), computer-mediated communication (CMC) and distributed artificial intelligence (DAI). The outcome was a three-level conceptual model of social context that combined top-down with bottom-up approaches in order to analyze social norms and activities. Barnard et al.'s (2000) "Systems of Interactors" theoretical frame work also drew upon several overlapping layers of macro theory and micro theory. Which level of theory is relevant depends on the nature of the problem being investigated.

While providing more comprehensive theories of HCI, these kinds of unified frameworks have proven difficult for other researchers to use in practice. It is much more unwieldy to juggle with multiple concepts, constraints and levels when analyzing a problem space and/or designing a system, compared with using a constrained framework that has far fewer interlinked concepts. While being attractive theories, it turns out that it is largely the authors, themselves, that have used them. Another case, perhaps, of the toothbrush syndrome.