3 Understanding Task Ecologies

Introduction

If we accept that acting on the world *is* cognition, then this ought to influence not only *how* we design but also *what* we design—or rather how we expect designed objects to be used. The dichotomy of form and function takes us only so far—and this is because the "interaction" between person and artifact does not occur in a vacuum but within an ecology shaped by the physical and the social environments in which artifacts are used. In chapter 1, the notion of a human-artifact-environment system was introduced. Much of the subsequent discussion focused on the first two elements, and I did not say much about environment and how this influences activity. Given that developments of Gibson's original theories have tended to come under the heading of ecological psychology (an academic discipline with its own society, journals, and conferences), "ecology" plays a key role in defining this field of enquiry. But this is jumping from the word environment to the word ecology without defining either.

For Gibson, "information" resolves uncertainty over which action to perform, and this makes sense only within an environment. Indeed, it is the role of the environment in his notion of "information" that separates his notion of uncertainty reduction from Shannon's. For Gibson, the environment is not an unstructured, amorphous mass of confounding and confusing features, but a collection of resources that are salient to action. In a very real sense, the environment (from Gibson's perspective) is analogous to the abstract notion of problem space used in chapter 2. For information-as-content, "salience" is a matter of assigning meaning to the symbols that the brain creates from features in the environment, which, in turn, requires storing symbol-meaning associations (hence, declarative knowledge in long-term memory). For information-as-context, "salience" relates to the mapping of environmental features to action (which makes no demand on symbol production or storage). At this point, the reader might cry "foul" and see this as a sneaky way of using representation to support cognition—one in which the mapping of features to actions is a form of representation. However, the point is that this "representation" is *not* a symbolic reconstruction of the environment (i.e., a "mental model") but a responsiveness to features in the environment that relate to action. To appreciate why this is the case, we need to define "ecology."

The Environment as an Ecology

Gibson distinguishes between the environment as the habitat in which a species lives and the "ecological niche" that supports the way of life of that species. In other words, the ecological niche of a species "refers more to how an animal lives than to where it lives."[1] So, two species of animal might occupy the same habitat, but, given their differing ways of life, might respond to the features of the habitat (in terms of what it offers for food or shelter) in different ways. "The natural environment offers many ways of life, and different animals have different ways of life. The niche implies a kind of animal, and the animal implies a kind of niche. Note the complementarity of the two."[2] The word "complementarity" is used to reflect the relationship between the animal and its environment. The implication here is that, whereas a species might respond to ecological information in one habitat, it might not respond in similar ways in another habitat. An obvious example of this concerns ways in which some (but not all) monkeys and apes use stones as tools, for example, to crack open nuts. In this usage, an ecological niche provides the resources for action. For much of the ecological psychology that builds on Gibson's theories, the ecological niche for humans consists primarily of the physical features that support specific actions (given specific abilities). However, he noted a distinction between the physical environment, which contains "information," and the social conventions that are superimposed on this ecological information. For the apes and monkeys mentioned above, the presence of stones or nuts only partially predict use of tools, and the culture (in terms of a tradition of tool use) of a particular troop would also contribute to the activity. Thus, the "ecological niche" consists of both physical features and cultural conventions. This is especially true for humans: "The material structure the human-environment offers should not, and cannot be ontologically separated from the social, technical and historical lives people lead."[3]

How Do Actions and Ecologies Interact?

While the proposal that the ecology consists of both physical features and social conventions has an intuitive appeal, it creates a problem in terms of the mechanisms by which actions relate to physical features. In a simple version, the physical features of the environment provide opportunities for action (so long as the animal can perform the action). This feels like behaviorism, in which the stimulus (of the physical feature) provokes a response (the action), as when Pavlov was able to condition dogs to salivate at the sound of a bell that they associated with food. Such unmediated pairing of stimulus and response misrepresents Gibson's ideas (and embodied cognition). But it also creates a problem for an approach to cognition that denies representation: If I am not saying that a feature directly causes an action, how does the relationship between feature and action arise, especially if the action will be influenced by social conventions?

As an initial response this question, I like Ingold's observation that "much if not all of what we are accustomed to call cultural variation in fact consists of variations of skills. By skills I do not mean techniques of the body, but the capabilities of action and perception of the whole organic being (indissolubly mind and body) situated in a richly structured environment."[4] Ingold used the term "co-respondence" to describe the ways in which an animal responds to its environment and the environment responds to the animal—in other words, the respondence is mutual, and so they "co" respond. This idea takes Gibson's notion of "complementarity," which is a rather passive stance, and turns it into the active interaction between animal and environment. The animal acquires the necessary skills through its ongoing relationship with the environment and processes of enculturation through which normative behaviors are learned (say, from observing its peers). Gibson called this the "education of attention,"[5] in that the ability to attend to relevant features is something that is learned and developed over time. In this way, normative actions become those that are performed by the majority of members of a given culture.

From these normative actions, the form of an artifact reflects social conventions, or majority ways of acting. However, mass production of artifacts satisfies a host of constraints in terms of materials and manufacturing processes to produce "one size fits all" designs which might impose rather than reflect normative action. It we step back to a time when artifacts were fashioned to suit local environments and local social conventions, then the argument becomes a little clearer. We noted in chapter 2 how the work of the wheelwright reflected the environment in which wheels would be used. As another example, a survey of shovel designs in 1930s Germany found over 12,000 alternatives.[6] While you might expect some variation for different tasks, and some variation for the use of different materials and cultural traditions across different regions, this is still a staggering number and suggests that the form that these shovels took was influenced by more than their basic functions, such as digging or lifting. The variations included the size and shape of the blade (reflecting the type of materials that they were required to work with) and the length of the handle (reflecting the physical activity involved in using the shovel and the normative way of performing this activity in that particular region). That is, "normativity" always relates to a concrete situation.[7] Acting normatively requires correspondence to artifacts (and other people) in *that* particular situation—for example, climbing *those* stairs in *that* building.

Responding to Ecologies

We noted previously that in order to manage the uncertainty in interactions, the elements of the human-artifact-environment system need to adapt to each other's behavior or to act is if their behavior is constrained. Taking the idea of "ecology" further, Newell[8] argues for three constraints on human behavior:

1. organism, which is defined by the size and shape of the actors, bodies and their capabilities to control their movements;
2. task, which is defined by the set of acceptable outcomes of action, in terms of how an intention is met but also in terms of how efficiently an action is performed;
3. environment, which is defined in terms of objects respond to general laws of physics, such as gravity or reflection.

From constraints 1 and 2, one can define combinations such that the perceptual capabilities of the organism can interact with the physics of the environment. The point is that embodiment is more than simply the possession of a body and has much more to do with the ongoing, reciprocal engagement between organism and environment in pursuit of task outcomes. Constraint 2 highlights that task outcomes are defined by the normative social conventions, in terms of what defines "acceptable." You could, for instance, eat from your knife, but it would not be acceptable in polite company. Of course, there are situations in which it might be acceptable—for example, if you are slicing an apple with a knife, it might not be a problem to raise the slice of apple on the knife to your mouth. This returns us to point that normativity is always related to a social situation.

Salience, Action, and Information-as-Context

An abiding question is how can features from the environment be defined as salient? In other words, what can we say about "meaning" if we are replacing content with context in our definition of information? In Gibson's terms, salience is directly tied to the action to perform and, crucially, does not require mediating activity (in the form of information-as-content). The broad premise of Gibson's work is that information exists as an array in the environment and that we respond to elements in this array according to the actions that we are performing. Cognition is thus less about information processing and more about response tuning.

Gibson's best-known account of this was in the form of optic flow experienced while moving through the environment. Gibson's initial thoughts on optic flow were inspired by discussions with pilots and his own experiences in flying and making training films for the US Air Force. As figure 3.1 illustrates, when landing an airplane, the pilot has the experience of heading toward a fixed point and the rest of the environment seems to be moving away from that point and "flowing" around the pilot.[9]

To account for this experience, Gibson proposed that patterns of light reaching the retina constitute an optic array that contains visual information from the environment (including the position of objects). The "flow" of this optic array results in a "textured gradient" that indicates speed (of self and moving objects in the environment) and relative distance. Responding

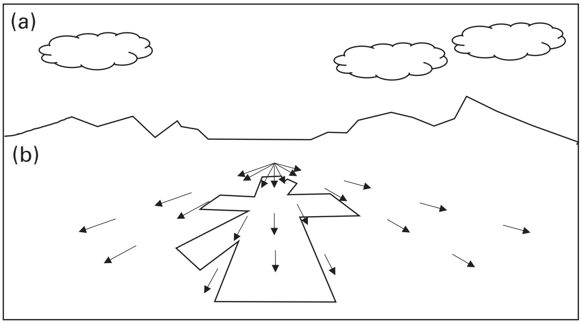

Figure 3.1
Optic array for a pilot landing an aircraft.

to changes in the optic array would be sufficient to support action, without the need for "information processing." Indeed, the very idea of optic flow is difficult to translate into information-as-content because it is not obvious which features of this flow can be represented symbolically. The labeling of features with symbols conjures up Garcia-Marquez's fictional village of Macondo where, during a sleeping sickness that robbed people of their memory, signs were hung on everything: "This is the cow. She must be milked every morning so that she will produce milk, and the milk must be boiled in order to be mixed with coffee to make coffee and milk."[10] Ironically, the sign conveys information-as-content with little clue as to how the actions (milking, boiling, mixing) ought to be performed.

The direct experience of the optic array does not require processing of information from the environment into a "mental model." Where there is uncertainty in visual information, this can be resolved by altering the experience of the optic flow, such as by moving the eyes or the head. The sampling of the optic array depends on the combination of the movement of the person and the content of the environment. Gibson proposed that people can "tune into" the optic flow in a specific environment in such a way as to allow their visual perceptual system to find resonance with the environment. From this, the environment is experienced directly, without the mediation of information processing. It is only a small step to associate elements in the optic flow with corresponding actions. For example, in the concept of "safe field of travel" (figure 3.2),[11] Gibson and Crooks suggested that the optic array could be considered in terms of regions around, for

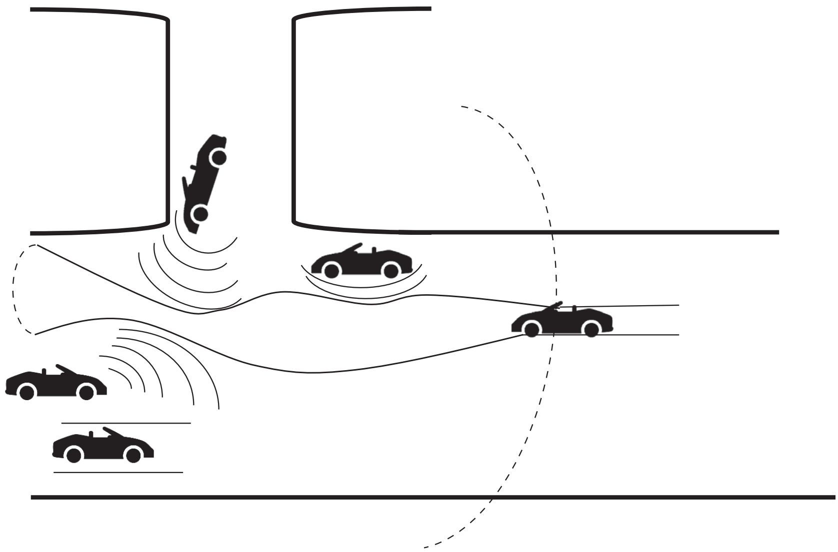

Figure 3.2
"Field of safe travel" for automobiles.

example, the automobile in which you are riding (either as driver or passenger) and those of other automobiles (or other objects on the road). Such "safe fields of travel" are defined primarily by relative position and speed in order to avoid collisions. This "safe field of travel" does not require the construction of a mental model of objects in the environment, but a continual sampling and adjustment relative to the optic array. In these examples, the environment is viewed as a complex pattern of "information" that changes in response to the actions that the person performs (or the actions of artifacts, people, and other elements). Activity in the environment is a continuous process of adapting to the available "information," either in terms of acting to avoid collisions, or to resolve ambiguity, or to modify the environment.One can conceptualize activity in the environment as a closed-loop (cybernetic) control system. Significantly, separating the elements of the system (i.e., human-automobile-environment) into constituent components is not feasible[12] and this points to another distinction between information processing and embodiment in that the borders between elements are permeable.

Brunswik's Lens Model

While Gibson's work is concerned predominantly with the relationship between movement and salient "information" (-as-context), a parallel theory explored the role of "information" (-as-content) in decision-making. Brunswik's lens model[13] of decision-making broadly involves two elements:

1. a set of cues in the environment that can be objectively associated with a given decision outcome, and
2. the selection of those cues according to a human decision.

Accordingly, cues (i.e., features of the environment) can be assigned a diagnostic value (i.e., a correlation or weighting in terms of the relation to the "correct" decision), which is the "true cue" validity. This can be contrasted with the cues that the decision maker chooses, or the "observed cue" validity. The performance of the decision maker can be evaluated by the correlation between these two forms of cue validity. Brunswik's proposal is that the environment contains cues that can be defined in terms of their salience to making a correct decision (figure 3.3). Contrary to "optic flow," Brunswik's model requires a series of steps by which the environment is perceived (in terms of an image on the retina), and this image forms the basis for probabilistic inferring of salience. In this way, there are mediating states through which "information" is translated, and this, although not directly related to symbolic information-processing approaches is part

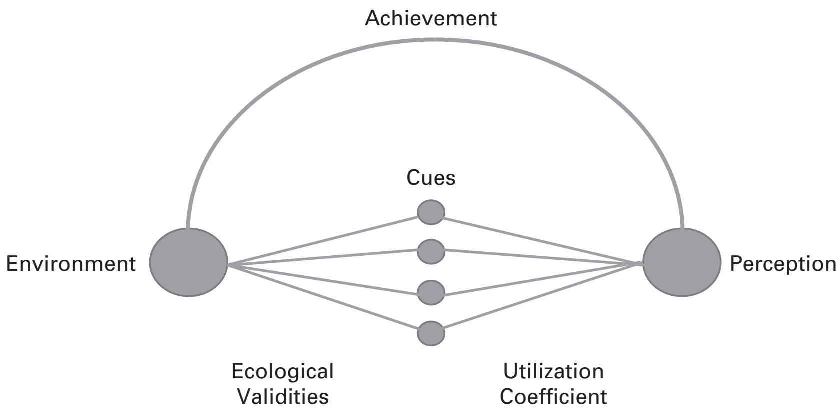

Figure 3.3
Illustration of Brunswik's lens model.

of the same family and hence, not an embodied cognition approach. However, there could be some benefit in considering how these might overlap.[14]

For Gibson, "information" arises from the one-to-one correspondence (complementarity) between an invariant feature in the environment and the action that the organism can perform (which, in turn, relates to the concept of "affordance," which will be considered in chapter 4). This means that there is no requirement for "translation" and that perception is direct. For Brunswik, the environment contains cues that are "context-free." In a study of stair-climbing, short and tall participants were asked to determine whether stairs could be climbed easily or not.[15] Plotting the data simply in terms of riser height (as a "context-free" metric) showed little correspondence across the two groups. However, plotting the data in terms of the ratio between riser height and leg length showed a neat correlation that was consistent across groups.[16] In other words, "information" is not context-free but rather involves an organism's specific frame of reference.

Both Gibson and Brunswik recognize that "perception" occurs in an animal-environment system and arises in terms of the capabilities of the animal to respond to its ecology.[17] Indeed, it is these aspects that influence Brunswik's emphasis on "ecological validity." For Brunswik, a "representative design" (of an experiment) would match the correlations between cues and outcome found in the real-world setting of the task. The reason why this is critical is that, for Brunswik, people learn to weight cues on the basis of successful outcomes (where, over time, they receive feedback on the accuracy of their decisions); indeed, this might allow the experienced person to discover relationships between cues (i.e., contiguous cues), which can allow predictions to be made accurately by sampling fewer cues (i.e., if cues are related, then one needs to sample only one rather than all of them). However, these correlations (between cues and between cues and outcomes) depend on the environment and on social convention; if the environment changes, the relations between cues and outcomes might change, with the result that performance will deteriorate; if social conventions change, then the definition of a successful outcome may change.

For Brunswik, the ability to apply strategy distinguishes experienced and inexperienced decision makers. As decision makers gain more experience of the environment, their ability to use the relationship between cues and outcome improves.[18] Indeed, the implication (in Gibsonian terms) is that the human will become tuned to specific cues (at least in familiar environments).

As they gain more exposure to these cues, combinations of cues could allow redundancy (which can improve reliability of decision outcome). If there are too many cues, then it becomes difficult to define salience. However, it is plausible to assume differential weighting of cues and that people will seek out those that are more salient to the task, and this is the basis for contemporary theories inspired by Brunswik, such as Gigerenzer's "take-the-best" heuristic.[19] In much of this work, the number of relations that are used is very small, amounting to two or three (with more than three being deemed "high").[20] Relating the number of cues to the discussion of multi-objective optimization, the finding that people tend to work with highly limited sets of relations (or cues) aligns neatly with the suggestion that design might focus on a limited number of objectives. In this, the definition of an "objective" (in terms of design) becomes a means of defining relations between cues, which implies that the weighting of cues (defining their salience) can have an impact on how these are sampled and selected.

Much of the research employing Brunswik's lens model focuses on static decision problems, rather than activity in dynamic environments. Consequently, the relationship between the "cue" (from the environment in which the decision is to be made, e.g., in terms of sets of "information-as-content" provided in an experiment) and the "perception" by which the cue's salience is defined becomes caught in a discrete, static task. For Gibson (and embodied cognition) the tasks are dynamic (like the "field-of-safe-travel" problems discussed earlier), and this makes it difficult to distinguish the cues in the ways in which Brunswik lens model assumes.

Exploring dynamic decision making, an experiment simulating an anti-air warfare coordinator,[21] performance was modeled using the lens model. The results suggested that "good" performance involved pattern recognition of salient information. Furthermore, training, feedback, and practice that allowed participants to refine the heuristics by which cue salience was defined improved performance. That is, performance did not involve appeal to mental models, but focused on perception and action. For Kirlik,[22] the mapping between the perception and environment becomes direct—that is, a one-to-one mapping—with experience. What is attractive about this is that it brings the possibility of rapprochement between the Gibsonian and Brunswikian approaches (although, of course, there remain clear and obvious differences in methodology and underlying theories) in that increasing the experience of an environment is likely to increase the consistency by which specific cues are deemed salient.

Recognition-Primed Decision-Making

One implication of the research using the Brunswik lens model is that people ought to attend to the most salient (rather than all) the available features, and that selection of features could be influenced by action (in a reciprocal manner to the action being influenced by the attended features). Recognition-primed decision-making (figure 3.4) interprets a situation in terms of "relevant cues." These are evaluated in terms of "plausible goals" and "expectancies" (the latter drawing on a mental model, or schema, that reflects experience of previous situations). In this model, when an action is defined as suitable, its outcome is simulated prior to implementation. However, work with expert decision makers tends to suggest that this simulation phase is rarely implemented.[23] This implies that decision processes (particularly in the highly dynamic, risky, and ambiguous situations that Klein and his colleagues study) is a matter of perception-action coupling. Indeed, an alternative version of this could be proposed in which there is no requirement for schema but in which the salience of cues (acquired through experience) are defined by the experts;[24] analogous to the description that the Brunswick lens model offers. Decision-making that is "automatic" and that operates through perception-action coupling not only provides an explanation of the fast "intuitive" decision-making that is required in emergency situations (and which people might find difficult to articulate because of the reliance on "tacit" knowledge) but also explains the role of expertise (gained from experience of many different types of situation) in shaping response.

As we saw in chapter 2, the way a person responds to a problem is highly dependent on the way in which the problem is presented and the physical actions that are possible in that context. In the words of Vallée-Tourangeau and colleagues, "A reasoner is embedded in a certain task environment that together configures a certain cognitive ecology within which certain cognitive abilities are manifested."[25] That is, the problem solver becomes immersed in the context of the problem, and this very immersion creates opportunities for acting and thinking. The implication is that being able to physically act in the "cognitive ecology" of the problem can be beneficial for problem-solving and,

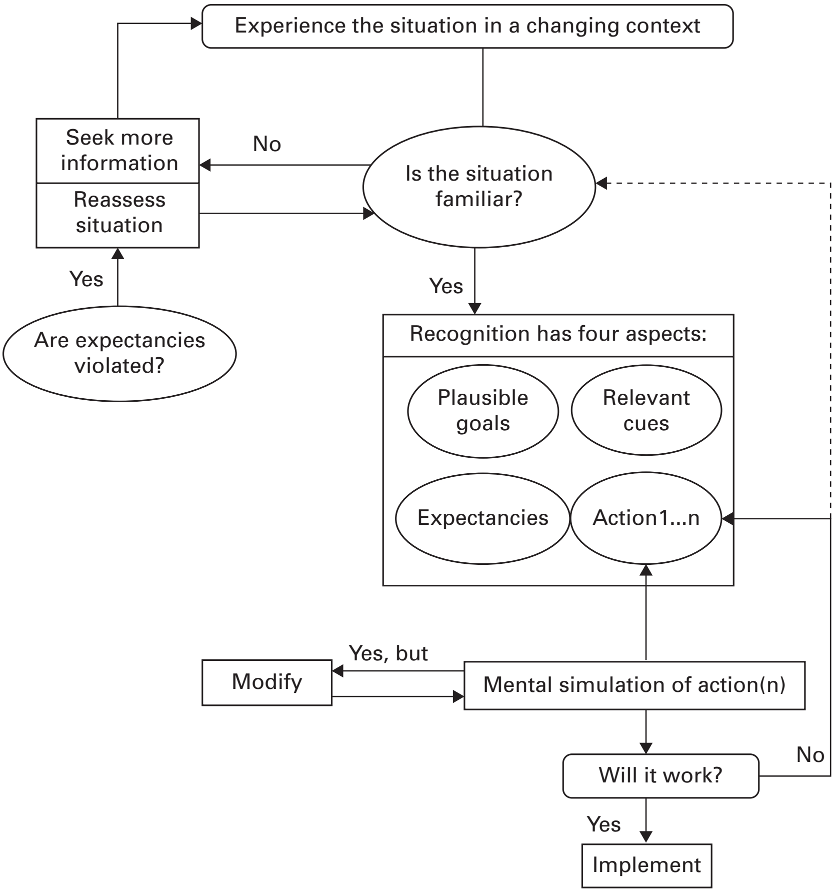

Figure 3.4
Recognition-primed decision-making.[26]

indeed, change the nature of the problem and the strategies used to solve it.[27] For now, the question to pursue concerns how "meaning" can be arrived at from an information-as-context perspective in terms of a "task ecology."

A "Task Ecology"

If radical embodied cognitive science is to be useful for design, it needs to reflect the ecology in which activity occurs and to do so in such a way as to provide rigorous, testable descriptions that could inform design decisions.

For the human-artifact-environment system, we ought to define activity not simply in terms of the single human using the single artifact, but in terms of the "context of use" (or ecology) of activity that carries a sense of "flow."[28]

From the initial discussions in this chapter, it should be clear that an environment can be described in functional terms—that is, relative to the goals and actions of a person, which constitutes the task ecology in which activity is performed. So, when driving a vehicle, the environment has features that support functions, such as the terrain (whether the ground is rough, muddy, slippery, whether there are hills, and so on), obstacles (other road users, pedestrians, objects on the road, and so on), and the path to follow (e.g., type of road, markings on the road, traffic signs, instructions from a navigation system, and so on). The salience of features is relative to the overall goal of the activity. This observation raises two key concepts. The first relates to the selection of salient features, and this could be considered in terms of their Gestalt (pattern) for the system—that is, the relationship between the whole "environment" and the "part" (pattern) that is salient to the goal. In this case, the goal becomes interpretable in terms of the system that performs the activity within the environment in which the activity is being performed. This relates directly to the idea of a human-artifact-environment system. The purpose of the system arises from a combination of the functions that it is configured to perform (some of which might interact with others) and the state of the environment in which it is operating. Each of these provides constraints on the system's operation. In this respect, the design should present only those combinations of features that are salient to the function at hand.

A second concept that arises from this notion of task ecology as a set of constraints is that the "meaning" of features (in the environment) becomes salient to activity. In this respect, the design of a user interface (see chapter 5), should present salient features in a manner that enables means-ends analysis. In this respect, a "task ecology" consists of the set of features in the environment that are relevant to a specific goal (performed by a system that is able to recognize and act upon those features), and so the design of a user interface would present salient features in a manner that allows direct perception of system state and definition of appropriate action. Here, only that part of the system relevant to the situation is presented, and this decomposition (from whole to part) becomes integral to design. The overarching framework of this notion of task ecology echoes the following quotation from Gibson:

> Things are components of other things. They would constitute a hierarchy except that this hierarchy is not categorical but full of transitions and overlaps. Hence, for the terrestrial environment, there is no special proper unit in terms of which it can be analysed for once and for all. There are no atomic units of the world considered as an environment. Instead, there are subordinate and superordinate units. The unit you choose for describing the environment depends on the level of the environment you choose to describe.[29]

The environment is a "space of possibilities," in which different constellations of features arise to afford different functions that can be realized through combinations of activity)[30]. The point is that the very of ideas of data or activity are seen, not as defining the environment, but as low-level components that constrain and support the functions of the system in that environment. So, how do we make sense of task ecologies as problem spaces.

Studying Task Ecologies

An ecology is a mesh of features defined in relation to different actions. Depending on the capabilities of the person, the experience of these relations will alter. This seems to me to be a reasonable formulation of what Merleau-Ponty called an "intentional object." which describes that collection of features in an object that correspond to a given intention for its user. I should note that Merleau-Ponty did not assume that the "intention" was some teleological impetus (that is, the "user" need not set out to achieve a definite outcome in "using" the object); rather, he saw the "intention" as becoming realized through the ensuing interaction.

In her analysis of how people interact with technology, Suchman[31] observed pairs of participants trying to use a complicated photocopier. Her study employed conversation analysis to reveal the ways in which people respond to the changes in state of the machine in relation to their own developing and emerging goals, as well as the ways that sense is made through their interactions with the device and each other. An insight from this work was that people respond to opportunities offered by the situation (or, at least, their interpretation of these opportunities) to choose which action to perform—and then, in retrospect, frame these choices as a plan. One can read the transcripts of the conversations as ways in which people struggle with the photocopier as an object that is "present-at-hand," rather than as something that facilitated their activity as "ready-to-hand."[32] Furthermore, her insistence on reflecting behavior in its situation (rather than only on the "tasks" performed by users of technology) creates a richer and more elaborate perspective to the one that Human-Computer Interaction (HCI) had typically employed before this work. However, by capturing specific instances (no matter how much detail the reports contain), ethnomethodology (in HCI and design in general) becomes "fact dependent."[33] This means that any generalizable principle (or design guideline or requirement) can be only an ad hoc response to the specific instance. So, the approaches offer no "guide to discovery" in which predictions can be made or conclusions empirically tested. In a sense, the collection of instances become like stamps to the philatelist. Such criticism can be leveled with equal force at other methods that rely on observing people as they work (and so is equally applicable to the ergonomics methods of task analysis). Addressing this criticism can involve either a sufficiently long period of immersion in the domain that is being observed or the development of a theory that can provide a "guide to discovery." In terms of the former, there is some concern with the manner in which HCI is currently practicing ethnomethodology, as expressed by two people who had been involved in the field from its early days. As Sharrock and Anderson write,

> We were part of the effort which opened up the promise of ethnographic fieldwork for software development. We still believe that it has a lot to offer. However, in the intervening years, we have watched as what can only be regarded as old-fashioned customer relationship management has masqueraded as ethnography. As the consultants have moved in, the canons of fieldwork rigour have been eroded. Now, it seems, any kind of conversation with any kind of user/manager/customer can be called ethnography.[34]

The emphasis on the richness of the context of use that Suchman called for can be seen in similar trends in philosophy, with phenomenological interviews,[35] and in the field of naturalistic decision-making, with its use of critical decision method interviews.[36] In each of these domains, the objective is not merely to elicit the sequence of tasks to perform or the information that is attended to, but to situate activity in the settings that gave rise to it. So, for example, a critical incident interview might encourage interviewees to reimagine the sounds, smells, environmental conditions, and so on in a situation before asking about how they dealt with an emergency. Similarly, Gallagher 's account of "expert performance" (in dance, sports, or music)[37] emphasizes *performative awareness*, by which the expert can attend to, and reflect upon, the movement of their body in ways that the less experienced person might not notice or appreciate. Having said that, the objective of the phenomenological interview and the critical decision method is not simply to reconstruct a specific situation but to obtain generalizable observations that can inform theory and design. That is, the aim of a phenomenological interview is not to capture idiosyncratic, personal experience, but to "capture the invariant structures of experience."[38] I wonder whether ethnomethodology can focus too much on the specifics of the situation, leaving the designer to translate the comments to a design concept. In this, the analyst becomes guilty of abrogating responsibility for translation of their findings. By "translation," I mean the abstraction of general points from specific accounts that then become formalized into design briefs, requirements, specifications, and the like. Because these approaches to analyzing technology in situ are fact-dependent, they have neither the aim nor the ability to serve as "guides to discovery." This means that someone needs to undertake the translation. Having the analyst integrated into the design team can dramatically reduce these translation problems. Certainly their presence can help in explaining the analysis, resolving confusion, and addressing contradictions. But the very act of "reducing" the analysis to a set of "user stories" or "persona" could become antithetical to the capturing of "real experience." Equally, because the analyst becomes a "proxy" for the end-users, there might be a danger that it is the experience of the analyst that is used to "stand in" for these users. This can be even more telling when the "analyst" is relying on previous experiences (as when subject-matter experts work in design teams). Continually reflecting the analysis (and its interpretation) back to the people who were studied or will be actually using the design can ease some of these problems.

Suchman's work set the tenor for a whole field of ethnomethodology in HCI. In this field, the environment is "meaningful primarily through the ways we interact with it."[39] At first glance, the emphasis here is on "the ways we interact" (and, indeed, Dourish's influential book is called *Where the Action Is*). However, ethnomethodology pays less attention to physical action (at least in terms of "action" as it could be described from fine-grained analysis of movement) and to the environment (at least in terms of the rich configuration of features that it offers) and more attention to the verbal account of the experience of action. This is not to say that movement or the environment is ignored; in many of these studies, scenes from video recordings of people interacting with technology are used as data and complement the verbal accounts. However, I have two concerns about this approach. The first is that the experience being studied comes partly from the words of the person performing the action and partly from the author's description of the video. This reiterates my point that the analysis is less about the action itself and more about the description of the action.

Aside from concerns over how "ethnography" might be practiced or how it might provide a "guide to discovery," my worry is deeper and concerns the focus on "meaning" rather than action per se. While Gibson regarded the environment in objective terms (that is he took a realist position in terms of ontology), he also wanted to be clear that the experience of the environment is shaped by individual capacities—that, the environment contains features that exist whether they are "experienced" or not. But when we consider the ways in which "experience" is collected, through ethnomethodology, as verbal reports, then "action" becomes the input to an interpreted account, rather than an activity in its own right. If we take the phenomenological position of Merleau-Ponty in as considered a manner as possible, then it would seem that asking people to provide verbal descriptions becomes almost the *least* effective way of capturing their experience. At best, such verbal descriptions are loose, informal, and incomplete accounts. The use of verbal description to derive meaning runs significant a risk of tipping from a phenomenological account of experience into one that is indistinguishable from the nominalist accounts of information processing from which it seeks to depart. This is because a verbal account, by definition, enforces a commitment to symbolic representation. So, this creates an impasse. Granted, ethnomethodology, the phenomenological interview, or the critical decision method take great pains to create a broad a view of the context. Granted, each of these approaches argues for the importance of context and (as has been argued throughout this book) acknowledges that action can only be considered in context. But there is a danger that an account that is heavily contingent on verbal descriptions (either of the participant or the observer) runs the risk of reporting the external, describable characteristics of the environment or the action and misses the underlying aspects of the context (in terms of the relations within the human-artifact-environment system). In a sense, my concern is that this approach could capture information-as-content (through its verbal description) and miss information-as-context (in the Gibsonian sense). In the words of Sharrock and Anderson quoted earlier, the risk is that ethnomethodology in HCI becomes "good old-fashioned customer relationship management."

Another way of seeing this is that the approach could reflect aspects of epistemology without fully appreciating the ontology. To do the latter would require a definition of the situation in terms of those features that are salient to the action.

The use of verbal reports as data[40] has been criticized in terms of the relation between such reports and information-processing. For instance, there is concern that people will describe only those aspects of their activity of which they are consciously aware (making it difficult to capture tacit knowledge or to fully reflect the impact of the environment on activity), or that people will describe only those aspects of activity that can be put into words (making it difficult to capture procedural knowledge), or that people might alter their activity to make it easier to describe. Such criticisms can be traced back to the challenges to introspection raised by the early experimental psychologists. In a sense, such criticisms point to a schism between those approaches that favor formally observed behavior in controlled settings and those that emphasize the importance of individual experience. This could be seen as a fight between laboratory-based approaches (central to much of the information-processing school of cognitive science) and ecological and experiential ones. But, from the perspective of embodied cognition presented in this book, the distinction is not quite as clear-cut.

Several of the studies that I have cited, in this chapter and in chapter 2, rely on conversation analysis to combine the words as they are spoken by people engaging in everyday activity. For some forms of analysis, video alone might be informative and instructive. For example, William Whyte created documentaries, such as *The Social Life of Small Urban Spaces*.[41] He provides a commentary on the activities of people, say in a public park, using different camera angles, to draw attention to aspects of the environment that we often miss or take for granted. The film allows us to "see the general pattern of behavior or sociocultural practice in this place, but when we zoom in we find a great variety of ways in which people engage with the various action possibilities the park offers."[42]

Combining video from, say, a head-mounted camera with verbal commentary (recorded either during the video capture or through interview afterward) provides a further means of capturing some aspects of the experience of the skillful coping of people as they undertake work in "real" settings as opposed to the artificial conditions of the laboratory.[43] The imagery (video or stills) gives context for the conversation analysis. In this way, the verbal reports takes precedence, with stills from the video presented to illustrate what is said. Notwithstanding the problems raised concerning verbal reports in the preceding discussion, it feels to me that relegating the video to illustrative purposes loses much of its benefit in capturing the dynamics of the interactions.

Richer analyses can be obtained by combining these words with stills from video recordings of the task performance.[44] This provides the opportunity to explore the temporal setting of the conversation as well as its physical and linguistic unfolding (figure 3.5).

From the perspective of "ecological validity" (as defined by Brunswik), the challenge is to adequately capture the cues that people use to inform their decision-making. Contemporary theorists such as Gigerenzer apply decision tasks in which cue selection is an essential feature of the experimental design. This allows consideration of the manner in which people select between the cues available to them. This could be done simply by hiding the cues until people select them—for example, by having information on a computer screen that becomes revealed only when the person clicks on it or through the use of eye-tracking. Combining these two approaches, my colleagues and I explored how people choose from a set of options in a simulated credit card fraud analysis task.[45] The layout of the information (in a grid on a computer screen) satisfies Brunswik's notion of ecological validity (in that the aim was to present the correlations between cues and outcome found in the "real-world" task).

Eye-tracking was used to explore strategies of information sampling when people copy a pattern presented on a screen. Participants tend to look at patterns at strategic moments.[46] For example, an initial glance might be to identify the color of the block, and a subsequent glance might be to determine its precise position in the pattern. Such a "minimal memory strategy" shows how people respond to task constraints in their sampling of the environment. Thus, the manner in which eye movements are performed is influenced by the type of task[47] Applying eye-tracking to working environments is more challenging but can reveal insights into the strategies that people use to sample their environment in order to make decisions.[48] Using eye-tracking in vivo, however, can be challenging to implement,[49] not least because the human eye is in constant motion ,and this means that where it pauses (fixates) need not be the point at which "information" is obtained (one could, for example, be gazing into space rather than looking

	0	.02	.04	.06	.08	.1	.12	.14	.16	.18	.20	.22
P5 ()								[speech not audible]				
P1 action	Sitting on edge of desk 1											
P2 (blue)												
P2 action		Marks map of UK on desk 1 [slide 5}					Stands			Walks to whiteboards		Writes on whiteboard
P3 (white)								Sunny Jim?				
P3 action	Looking at statement of harbor master log and harbor								Walks to whiteboard			
P4 (grey)	Yeah		Yeah		Yeah yeah yeah				Sunny Jim is the name of the boat.			
P4 action	Looking at statement of harbor master log and harbor map										Walks to whiteboard	
P5 (red)	…this boat (0.2)…ah…that docked here…which is really close to the car park (0.1) and there's another…ah…											
P5 action	Looking at statement of harbor master log and harbor map											

Figure 3.5

Communication between participants in a sense-making exercise.

directly at something). However, like the approaches that combine conversation analysis with video recording, eye-tracking provides a means by which investigation can attend to both the properties of the environment (as these impact on the actions being performed) and the dynamics of the activity.

In a study of the micro-materiality of the handling of surgical instruments, Heath and colleagues noted how little talk occurs. Consequently, their analysis uses "a series of images accompanied by brief descriptions in relation to the timing of the particular activities. . . . In fragment 1 for example, the activities in question commence 16 seconds into the fragment and last for just over 3 seconds."[50] Regarding the activity of a scrub nurse passing a dilator to the surgeon, they note that "the scrub nurse clasps the dilator at both ends and passes it horizontally to the surgeon. The surgeon is able to grasp the dilator with his thumb and forefinger in the center of the rod. Without adjusting or repositioning his hand, he immediately inserts the head of the dilator into the patient's oesophagus."[51] This evocative description allows them to discuss not only the physical actions that are performed but also the knowledge that scrub nurse and surgeon need to have in order to collaborate, with each making sense of the other's actions within the social milieu of the operating theater. Here, the task ecology is the social and physical environment, rich in the culture of surgical practice. If we are to understand a given task ecology, then it is studies like this that will provide the level of detail required. However, while the description (of the scrub nurse and the surgeon) is detailed, it lacks the nuance and richness that could be obtained from the original video (or from data that would fully describe the performance of the activity).

The focus on "materiality" in these studies aligns neatly with the concepts in this book. In this sense, materiality concerns the questions surrounding the physical properties of artifacts and the impact of these on how they are used.[52] Recent work by Dourish[53] extends the concept of materiality to the ways in which digital technology represents and allows interpretation of information. Notwithstanding my discussion of the different uses of the terms of information (for Dourish's analysis the focus is on information-as-content), his work emphasizes the need to appreciate how interpretations as intended by the designers or managers of technology might differ from those of its users. To return to the example of handling surgical instruments describe by Heath above, each of the actors in this activity might regard the same artifact as having different meanings, which then compel different opportunities for action. The physical appearance of the artifact and the environment in which it is used define, in a sense, the ontology of the actors. Their interpretation of the artifact forms their epistemology, which reflects the ecological niches in which each of them defines the appropriateness of action. The implication is that the mere physical appearance of the artifact is not an imperative to action; one cannot simply assume that form dictates a single function. Nor, as I argue in the next chapter, does it make sense to speak of "affordance" being the property of the object. Rather, the materiality of the artifact is captured by both its form and meaning to the person using it for the purpose that person intends.

To this end, adding sensors to the person or the artifacts they use[54] allows fine-grained analysis of activity. While much of the work on human activity recognition and analysis focuses on identifying specific, discrete actions, it makes more sense to understand activity in terms of the manner in which movements balance between consistency and variability, as, for example, in terms of Bernstein's notion of dexterity. Rather than asking "what" action is being performed, it is more useful to ask "how" actions are being performed, and this involves the application of dynamic systems approaches to the study of human activity. In such approaches, the analysis considers the temporal variation in the control and organization of patterns of movement, often using metrics related to entropy. Radical embodied cognitive science makes extensive use of these metrics and concepts, and this will be explored in chapter 7. In the next chapter, I focus on the ways in which perception-action coupling in the use of physical artifacts has been explored using the concept of affordance.

Embodying Design

An Applied Science of Radical Embodied Cognition

By: Christopher Baber

Citation:

Embodying Design: An Applied Science of Radical Embodied Cognition
By: Christopher Baber
DOI: 10.7551/mitpress/12419.001.0001
ISBN (electronic): 9780262369886
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Baber, Christopher, 1964– author.
Title: Embodying design : an applied science of radical embodied cognition / Christopher Baber.
Description: Cambridge, Massachusetts : The MIT Press, [2021] | Includes bibliographical references and index.
Identifiers: LCCN 2021033926 | ISBN 9780262543781 (paperback)
Subjects: LCSH: Expert systems (Computer science) | Human-machine systems. | Thought and thinking. | Artificial intelligence.
Classification: LCC QA76.76.E95 B22 2021 | DDC 006.3/3—dc23
LC record available at https://lccn.loc.gov/2021033926