John Benjamins Publishing Company

This is a contribution from *Pragmatics & Cognition 14:2*

Distributed cognition

A methodological note*

David Kirsh
University of California San Diego

Humans are closely coupled with their environments. They rely on being 'embedded' to help coordinate the use of their internal cognitive resources with external tools and resources. Consequently, everyday cognition, even cognition in the absence of others, may be viewed as partially distributed. As cognitive scientists our job is to discover and explain the principles governing this distribution: principles of coordination, externalization, and interaction. As designers our job is to use these principles, especially if they can be converted to metrics, in order to invent and evaluate candidate designs. After discussing a few principles of interaction and embedding I discuss the usefulness of a range of metrics derived from economics, computational complexity, and psychology.

Keywords: epistemic actions, coordination, interaction, design theory, ethnography

1. Introduction

People interact with artifacts, technologies, surfaces and other people in ways that overwhelm our current formalisms. Dynamical systems theory, game theory, economic theory, the formal theory of distributed computation, all fall short of providing satisfying explanations. Formalisms based on abstractions such as task environments fare no better. They are known to be flawed in making unreasonable assumptions about how rational people are, about how the environment of action can be characterized as a fixed set of choice points with fixed option sets rather than a place where people dynamically engage the world interactively. Look at any realistic environment — offices, cafés, computer 'desktops' — all these defy reduction to layerings or intersections of task environments. How are we to proceed in our search for a scientific study?

Pragmatics & Cognition 14:2 (2006), **249–262**.
ISSN 0929–0907 / E-ISSN 1569–9943

Despite the importance of taking a system view, my own approach is to move from the individual to the group. If we can understand how individual people engage their environments, how they appropriate artifacts, how they rely on material aspects of their activity space to help them stay in control, to manage thought, perception and choice, then perhaps we can begin to put these individuals together into larger socio-technical systems. It is not easy to predict how individual behavior changes once people are put together. Coordinating forces, at the system level, help constrain individuals; but in the end people still act locally. Distributed cognition remains a challenge precisely because it is so hard to balance the reality of local choice with system constraint. We need better theories of how people are embedded in the world as well as better theories of how the world and the larger systems we are part of coordinate action.

The following six assumptions have guided my own research in looking for principles of interaction that might fit this bottom-up top-down model.

1. We act locally and are closely coupled to our local environments.
2. We externalize thought and intention to harness external sources of cognitive power.
3. Economic metrics have a place in evaluating distributed systems, but they must be complemented with studies of computational complexity, descriptive complexity and new metrics yet to be defined.
4. The best metrics apply at many levels of analysis, from the system level where our concern is with the goodness of a system's design to the level of individual artifacts, where our concern is with the goodness of the design of the artifacts individuals interact with.
5. Coordination is the glue of distributed cognition and it occurs at all levels of analysis.
6. History matters.

My objective in this note is to introduce each idea in a personal way, as an opportunity to open a dialogue with designers interested in shaping interactivity, which is one of the topics of this special issue.

2. Closely coupled

Let us say that two entities are closely coupled if they reciprocally interact: changes in one cause changes in the other, and the process goes back and forth in such a way that we cannot explain the state trajectory of the one without looking at the state trajectory of the other. When a person writes on paper, the two form a reciprocal system. The person causes paper changes, paper changes partially cause person changes. This reciprocal interaction allows the person to find expressions, to represent and explore ideas using the persistent state of the paper that would otherwise be impossible. There is a dynamic between the two.

One of my first appreciations of the practical value of seeing cognition as distributed, the outcome of a close coupling with external processes, came from studying Tetris (see Figure 1). In work on Tetris, Paul Maglio and I found that players had to be closely coupled to the game in both an epistemic and pragmatic way to play well (Kirsh and Maglio 1994).

The pragmatic side of coupling is uncontentious. In video games, where what happens next depends substantially on what a player does a moment before, interaction is intense, high frequency and the two, player and game, are closely coupled. They form a distributed system whose trajectory cannot be understood without explaining the role each plays in driving the next state.

In Tetris, however, we found that it was not possible to *understand* the actions players took if we looked just at their pragmatic goals in playing. They seemed to be coupled to the system epistemically too, often performing actions not for pragmatic advantage but rather for epistemic reasons. For instance, given a choice between performing a mental action of rotating a mental Tetris piece or a physical action of rotating a physical Tetris piece they often chose the physical action. The net epistemic effect of knowing what a piece would look

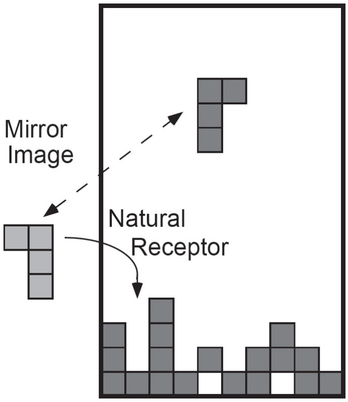


Figure 1. In Tetris the goal of play is to relentlessly fill gaps on the bottom layers so as to complete rows. The game ends when the board clogs up and no more pieces can enter. Because there are mirror pieces, and the choice of where to place a piece is strategic, Tetris players need many hours of practice to become expert. We found that players often rotate pieces to speed up their identification process, especially among pieces with mirror counterparts.

like rotated 90° or 180° would be the same. But physical rotation was faster than mental rotation so the same epistemic state could be reached faster if done physically. Physical rotation would have been pointless if it were not coupled so closely in time with thinking.

Our analysis let us say that these 'extra' moves paid enough in *epistemic* benefits to cover their pragmatic costs. In Tetris we were able to use speed accuracy of judgments and the impact of those judgments on game score to develop a metric for evaluation. Thus benefits, measured in terms of increased speed of piece recognition, improved choice of placement location, and improved control and visual judgment, more than outweighed costs, measured as the number of steps off the most direct path and the time to recover.

To mark this distinction in the role of actions we called intentional movements taken to bring a subject physically closer to its external goals *pragmatic actions* and those intended to simplify computation, reduce error or increase precision, *epistemic actions*. The distinction was not hard and fast. Certain pragmatic actions could also serve epistemic ends. And epistemic actions could be seen as pragmatic with respect to advancing epistemic ends when these were the external goal. But the point was that it is not always obvious why someone does something. We may think they are very goal directed in a simple sense. But without a better theory of exactly what is going on in their head and their world we might easily approach a task and mischaracterize the actions being taken. Actions might be either pragmatic or epistemic. Therefore the metric to evaluate action strategies, artifacts and technologies, may have to have pragmatic and epistemic components and be completely non-apparent.

Epistemic actions are everywhere. Some are connected with uncovering information. We move our hands close to something hot before grasping it. In woodworking the adage is measure twice cut once. Some epistemic actions compensate for sensory limitations. A trick most of us learned as kids was to squint our eyes or make a little hole with our fingers to look through in order to see distant items more clearly. Others are interactive strategies for externalizing. Say something quietly out loud to make sure it is coming out right — this serves to confirm. Some serve as reminders. Thoughtfully place one's keys in front of the door or in one's shoes to save relying on prospective memory alone. There are many other forms of epistemic actions. All have personal payoffs and depend on interaction with the environment.

The way to discover epistemic actions, as with most close interactions actions, is to record activity by video or computer and then ethnographically analyze it. We did this with Tetris and we also data mined our records of agent-environment activity. But in contexts where there is no predetermined metric

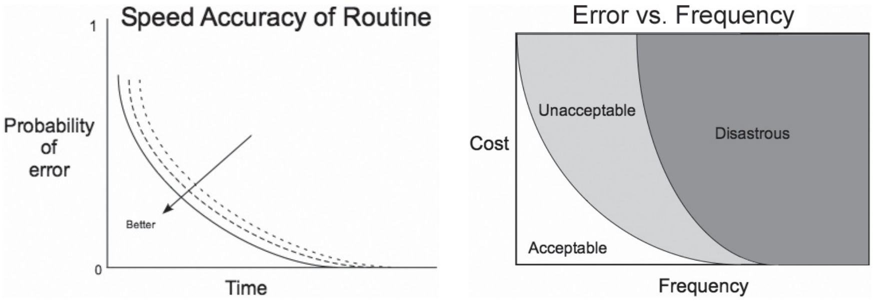
— in activities other than playing video games, for instance — it is hard to measure cognitive and physical costs and benefits.

3. Cost structure

My own efforts to develop metrics for evaluating the costs and benefits of actions, resources, and environments draw on ideas from complexity theory and economics. Consider economic measures first.

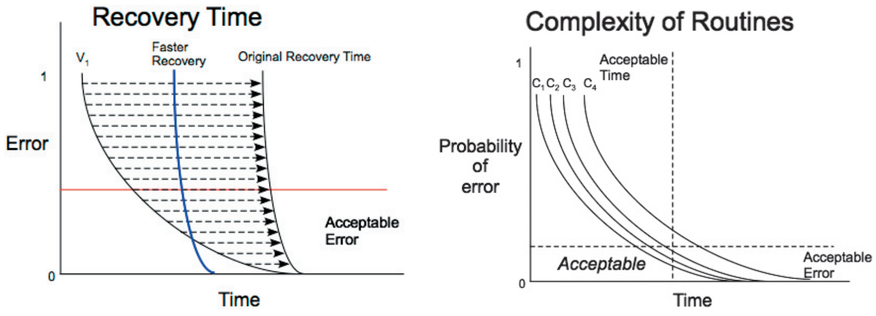
In recent ethnographic work at a local café I will call Jolt's, I have been studying the subsystem for taking customer orders and communicating them to the baristas who make the drinks. To develop metrics I thought it would be helpful to contrast the method at Jolt's with the method at another café which I will call Buck's. Buck's still relies on a paper cup that is supposed to be written on; Jolt's uses an IBM ordering system, with an overhead monitor displaying the queue of orders. Both are excellent mechanisms for coordinating a complex system of parts and processes, but each has strengths and weaknesses. The goal of the study was to develop objective ways of measuring some of these strengths and weaknesses.

I considered six metrics: speed accuracy, error type vs. frequency, error recovery rate, variance, learnability, and drink complexity. See Figures 2, 3, and 4.



a. b.

Figure 2 a. A speed accuracy graph shows the average time it takes skilled users to place an order vs. the probability that their entry is wrong. As users get better at their job we assume that they reach a maximum level of performance that is close to some ideal level given the technology, artifacts and procedures they are operating with. A better techno-social system ought to translate into better speed accuracy curves.
b. Because errors come in different flavors with different consequences, they need to be categorized, assessed for frequency, and their probable impact evaluated.

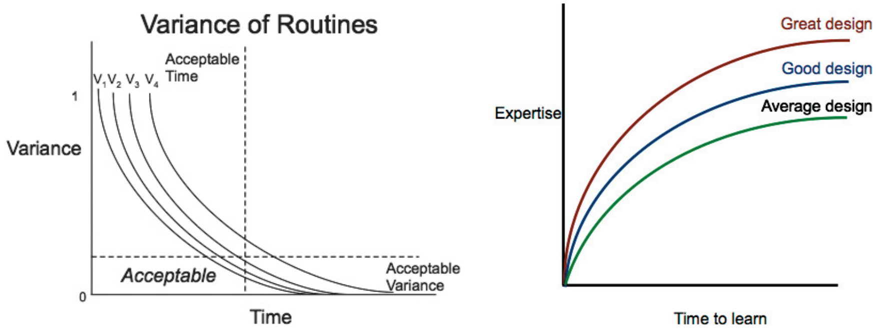


a. b.

Figure 3 a. A natural way to measure the value of the resilience of a system is to observe the time it takes the system to recover from an error, an interruption, or an error in process, and use that measure to compare techno-social systems.
b. At modern cafés drink complexity has risen dramatically. An excellent technology such as the Buck's cup need not increase production speed if it permits staff to produce drinks of greater complexity in acceptable time and quality.

Anyone interested in the evolution and spread of routines in cafés or other organizations would do well to compare routines using graphs and measures as in Figures 2 and 3. But there are a few things to appreciate about these sorts of economic analyses.

First, each graph measures a different parameter. This raises the question: how do we decide if it is good to introduce an artifact that improves speed accuracy but is worse under some other measures such as being harder to learn?



a. b.

Figure 4 a. A basic attribute of a good design is that it reduces the variance of output. The holy grail of quality control is uniformity.
b. Both Jolt's and Buck's have relatively high turnover in staff. Other things being equal, one system is better than another if new staff can more quickly learn the ordering and espresso making system, reaching the same performance level sooner.

If there is no independent measure of value for comparing the relative goodness of different parameters, then we must rely on a Pareto style assessment where all that can be said is: Env_1 is better than Env_2 only if Env_1 is at least as good as Env_2 in every parameter and better in at least one. This hardly addresses the reality of tradeoffs in design.

Second, these sort of quantitative abstractions don't explain *why* Env_1 is better. We have no mechanism. Thus in saying that the Buck's cup yields a better speed accuracy and error recovery rate than the Jolt's monitor system, we have *not* said what about the cup and monitor system make it so. For that we need a microanalysis of the type distributed cognition has become well known for (see, for example, Hutchins 1995). But ethnographic studies tend not to lead to generalizations about mechanisms — which seems necessary for a science of design.

Third, the strength or weakness of a technology will not be apparent unless we also include a careful account of what Vicente (2004) has called its 'human technology'. An example of human technology is the protocol required when taking an order. In Jolt's, as in Buck's, cashiers are required to repeat the drink specification to the customer. This recruits the customer's help in verifying the entry and has as a side effect that an alert barista, if nearby, can anticipate the drink. This formal practice reduces error. But practice is a complex phenomenon in itself. Metrics that are used to measure how good a technology is leave unexplained the interdependence of technology and practice. Performance is assumed to vary with skill (or learnability) but nothing is said about the nature of this skill.

To quantify practices I have tried, at times, to approach them as computational systems — routines really — that involve procedures, memory, control of attention, skill and reasoning. If a practice can be interpreted computationally then we can measure its complexity. How fast do memory requirements, the number of steps involved, and the number of items to monitor, rise as the size of the task is increased? How do cashiers cope with very large orders, or with two people ordering at once, or with one person giving two orders? What is the complexity of the momentary computational problems they face? Can they reduce these problems by doing things like making annotations as they work, by laying out physical objects to encode hints or cues? Do they have a way of talking to clients that lets them enter an order in a step by step manner so that they don't have to remember as much at each moment, or so that they are less likely to make an error while dual tasking?

I undertook a variety of these small computational analyses (Kirsh 1995, 1996). See Figure 5 for another example of how we can measure the effect of

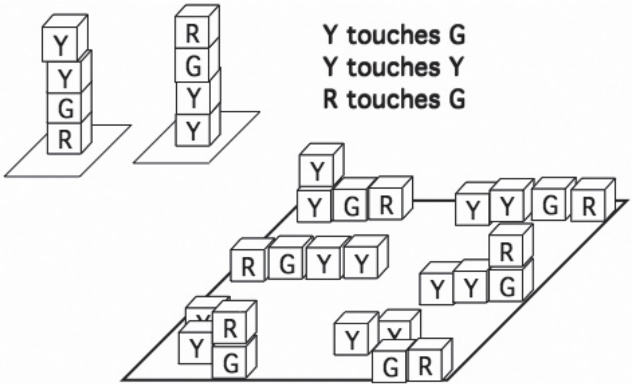


Figure 5. Just changing the size of the mat on which a problem has to be solved can change its complexity. In this case the blocks must be arranged so that the following constraint is met: Y touches G, Y touches Y, R touches G. Solving this problem on a small mat restricts the solution to the two towers shown. Working on a large mat however increases the solution space because now there are two dimensions along which the constraints can be met.

simple environmental manipulations on the complexity of a task. Tasks can be inventively changed to change their complexity profile. The question is: do participants reshape their tasks and environments this way? At this point ethnography becomes indispensable. Of the many things that participants might do to save cognitive effort, which are the ones they actually do and why? For instance, if a computational analysis suggests that a given practice should be difficult to master, and if ethnography shows that it is nonetheless prevalent, then we now have a specific question to ask: what is it about the way people use their resources that makes this problem easier than we assumed? Or is it done for historical reasons, because it is the way people always have?

4. Cognitive efficiency of design

My argument so far has been that, although economic metrics are helpful and necessary to analyze techno-social systems, it is necessary to observe the way people do things in ever more detail and tie these actions to cognition and computational analyses to get at the mechanisms of distributed cognition. The same type of analyses can be used to begin to measure and explain the effectiveness and efficiency of an artifact's design.

To develop deeper intuitions about the effectiveness and cognitive efficiency of designs — in particular visual designs — examine the following two figures. Why is Figure 7 better than Figure 6?

To answer this sort of question we need to develop cognitively motivated principles of design. What drives the way people interact with artifacts? In Figures 6 and 7 the point of the artifact is to structure and present forced choices. Any design that makes the structure and set of choices easier to appreciate, more visible, is a better design. The visibility principle at play here is 'What goes together semantically should go together visually'.

In Figure 6 the space between choices, labels and categories is so *ad hoc* that it is not easy to scan the figure and quickly say without looking at the labels what group of radio buttons one is choosing between. There is no clustering, no use of Gestalt principles of grouping to reflect in the layout the options which belong together. It is true that alignment is used to indicate a connection between category (e.g., Family) and options (e.g., Courier). But if we take as measures of efficiency the time it takes to scan this image and formulate a simple plan, the time it takes to determine if one has completed the entire form, and the time it takes to decide how much is left to be done, all will show that Figure 6 is less efficient than Figure 7.

If we were to give metrics for comparing two designs they would look much like our economic measures. For example, since subjects must decide if they have completed their task we can create a simple decision task and experimentally compare the speed accuracy curves for Figures 6 vs. 7. We can measure what types of errors are made and how frequently. Similarly we can

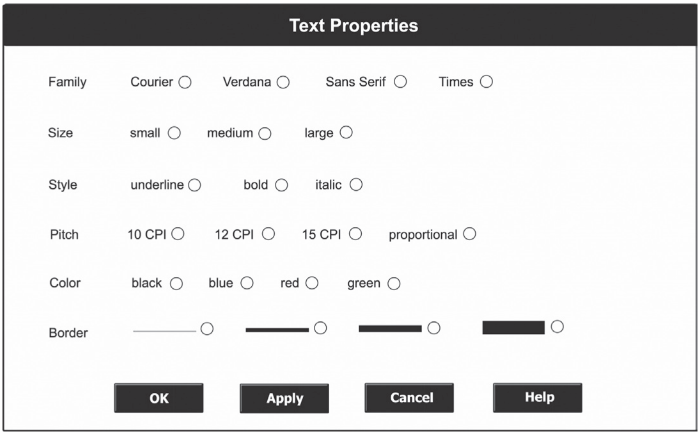


Figure 6. Here is a dialogue box for setting the text properties of a style in a word processor like MS Word in the 80's and the days of Microsoft DOS.

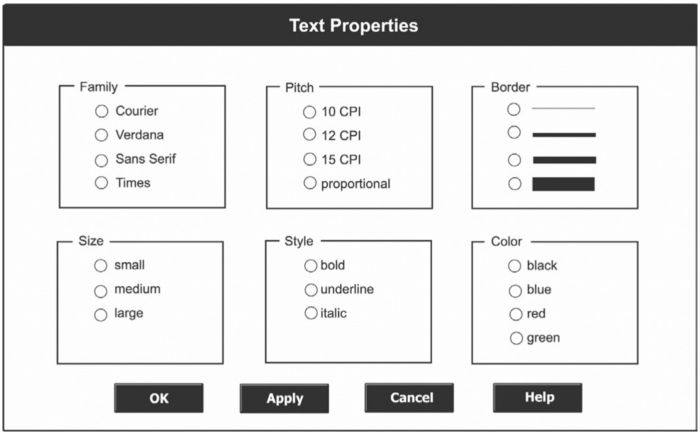


Figure 7. The same labels have now been distributed more effectively, using borders and positioning to simplify the perception of groups. The speed accuracy measures of performance on this layout trump those of Figure 6.

explore cognitive costs through computational analysis: suppose the problem were increased in size (more parameters to set). Are there more efficient algorithms available for one of the layouts than the other?

5. Coordination

The study of distributed cognition is very substantially the study of the variety and subtlety of coordination. One key question which the theory of distributed cognition endeavors to answer is how the elements and components in a distributed system — people, tools, forms, equipment, maps and less obvious resources — can be coordinated well enough to allow the system to accomplish its tasks. Even coordinating mechanisms as simple as clocks or paper clips can make the difference between a successful system and an unsuccessful one. Clearly we would like methods and measures for systematically exploring coordination.

Part of my own pleasure in studying humans comes from discovering the remarkably diverse ways we coordinate activity. So far I have briefly mentioned how individuals coordinate different parts of cognition by exploiting the environment. People manipulate local conditions to stay in control, to perform faster and more effectively. They annotate to cue response and reduce prospective memory, they line up items to make them easier to scan, to notice outliers and so on.

These same sorts of principles are exploited by good designers when they make artifacts that make our life easier. But all these examples of coordination are local; they affect local choice. In distributed systems the success of the whole depends equally on all these acts of local choice adding up, working together to move the system closer toward system goals. In a café it's about the beverages not the baristas. The system works well if clients get a satisfying drink, the one they ordered, quickly, at the right price, and well made. For that to happen, everyone's roles must fit.

Decisions about the roles people will play in a system, like the decisions about the artifacts, physical layout, routines and local goals seem to be on a different level than local choice. They have a lot to do with more global considerations about how everything fits together. Assembly lines have to be planned and laid out. Orchestral conductors need to make global choices about tempo and expressiveness. If these are not good then everyone can play their part perfectly but the overall product will be imperfect. Even good cooks, using good ingredients produce bad food if their recipe is wanting.

To study coordination at this level requires methods we have not discussed: modeling and simulation, scheduling theory, and others. Economic and computational models are of use in showing us the impact of different global coordinating mechanisms when we have actual systems to compare. If we do not have living versions of different systems of coordination, how can we predict the value of re-engineering a process? Only by modeling and simulating can we study the temporal effects of such things as changing the time and destination of resources, or the impact of changing the connectivity, reliability or speed of communication, or the pattern of messaging. Only through simulation can we begin to see how one participant's local activity in his own activity space can have side effects on neighboring or intersecting activity spaces and so produce a cascade of side effects.

For these reasons, modeling and simulation are necessary for exploring coordination and the robustness of a distributed system. But I'd like to close with two cautionary points concerning simulation and models. The first concerns the creativity which well motivated participants show in compensating for bad design or for torpedoing good design. The second concerns the importance of history.

In our Jolt's study baristas work in such tight quarters that their activity spaces physically overlap. This means that one person may put an item in the area another was intentionally keeping clear. With practice, baristas who share space get to know each others' work patterns and the two develop a new and better dynamic of coordination. This learning is idiosyncratic and hard to simulate.

Even more spectacular is the way people discover or learn to compensate for their own and their team's limitations. People with advanced Parkinson's disease, for example, lose the ability to control walking because of corruption in neural processing of proprioceptive input. Yet, if they walk on floors that have large stripes or checkerboard tiling they can compensate for the loss of proprioception by using the rhythmic input the visual system provides. How could such successes be predicted by models?

History is important because coordination in ongoing systems is almost always history dependent. To appreciate how hard adding history makes the problem of coordination, imagine that we set out to model and then simulate a distributed system in which agents rely on a clock to coordinate timing. Under reasonable assumptions we may be able to show that without a clock timing would be unacceptably bad. Great result! But where did our reasonable assumptions come from? Presumably from an idealization about the way the system in question operates right now. Yet if we have learned anything from looking at the complexity of systems, it is that evolution can find multiple paths to the same goal. For a large class of systems, including our target system, a differently designed system which relies on, say, conveyor belts moving at fixed speed covering a fixed distance can be temporally coordinated as well as any system with a clock. It depends on what needs to be where and when. This diversity of solutions highlights the need to stay close to the facts. We can never understand the elements driving the coordination of a natural distributed system if we suppose that the system, its setup, its timing, its rules and culture of operation, are devoid of history. Parts have been adapted at every level, and the form they are in now are a partial function of the form they were in before. If it were not so hard to know the aspects of a system that transfer well to the real world, business models would be more successful.

The upshot is that as designers we must always work from the present, mindful of the inertia of users. If we create a design that is too distant from current activities, however cognitively efficient we think it is, users will either not adopt it — so it is de facto ineffective — or users will co-opt it for their own purposes. The gulf between the theories we have and the designs we need remains wide.

6. Conclusion

In the search for a comprehensive view of how humans structure and interact with their environments, distributed cognition adds a meaningful constraint on the final story. What makes an environment a good place in which to work? Cognitive engineering would be considerably easier if we had a principled answer to this question. Designers could then consult the book of cognitive and interactive principles and, like civil engineers designing bridges, they could creatively apply principles to build environments that make life easier. At present, no such book exists. And the prospects of ever writing one depends on our understanding of a collection of issues broadly related to interaction and distributed cognition which we are just beginning to appreciate. In this paper, I have presented a few of these issues. Chiefly my focus has been on methods of measurement: economic, computational and conformity to principle. The complexity of real-world distributed cognition systems is so great and so dependent on details of history, personal expertise and physical layout that our models and metrics are bound to be inadequate. But we also are at a point in science where new mathematics, new modeling possibilities, and more extended videographic analyses will let us enter into the next level of measurement and observation. It is a good time to be a cognitive scientist.

Note

* It is a pleasure to thank Stevan Harnad and Sam Harvey.

References

Hutchins, E. 1995. *Cognition in the Wild*. Cambridge, MA: The MIT Press.

Kirsh, D. 1995. “The Intelligent Use of Space”. *Artificial Intelligence* 73(1–2): 31–68.

Kirsh, D. 1996. “Adapting the environment instead of oneself”. *Adaptive Behavior* 4(3–4): 415–452.

Kirsh, D. 1999. “Distributed cognition, coordination and environment design”. In *Proceedings of the European conference on Cognitive Science*, 1–11.

Kirsh, D. 2001. “The context of work”. *Human Computer Interaction* 16(2–4): 305–322.

Kirsh, D. 2004. “Metacognition, distributed cognition and visual design”. In P. Gardinfors and P. Johansson (eds), *Cognition, Education and Communication Technology*. Hillsdale, NJ: Lawrence Erlbaum, 147–179.

Kirsh, D. and Maglio, P. 1994. “On distinguishing epistemic from pragmatic actions”. *Cognitive Science* 18: 513–549.

McNeill D. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: The University of Chicago Press.

Vicente, K. 2004. *The Human Factor: Revolutionizing the Way People Live with Technology*. London: Taylor and Francis.

Author's address

David Kirsh
Department of Cognitive Science
University of California San Diego
La Jolla, CA 92093–0515
U.S.A.

kirsh@ucsd.edu
http://adrenaline.ucsd.edu/kirsh

About the author

David Kirsh is Professor in the Cognitive Science Department at the University of California San Diego, where he has been since its inception in 1989. His D.Phil is from Oxford University in Philosophy and he held a post doc and later was research faculty at MIT in the Artificial Intelligence Lab. His interests are broad but recently concentrate on understanding principles for designing interactive environments.