Intellectica, 2000/1, 30, pp. 19-51

David KIRSH*

A Few Thoughts on Cognitive Overload

This article addresses three main questions: What causes cognitive overload in the workplace? What analytical framework should be used to understand how agents interact with their work environments? How can environments be restructured to improve the cognitive workflow of agents? Four primary causes of overload are identified: too much information supply, too much information demand, constant multi-tasking and interruption, and inadequate workplace infrastructure to help reduce the need for planning, monitoring, reminding, reclassifying information, etc... The first step in reducing the cognitive impact of these causes is to enrich classical frameworks for understanding work environments, such as Newell and Simon's notion of a task environment, by recognizing that our actual workplace is a superposition of many specific environments – activity spaces – which we slip between. Each has its own cost structure arising from the tools and resources available, including the cognitive strategies and interpretational frameworks of individual agents. These cognitive factors are significant, affecting how easy or difficult it is to perform an action, such as finding a specific paper in a "messy" desk. A few simple examples show how work environments can be redesigned and how restructuring can alter the cost structure of activity spaces.

Keywords: cognitive workflow, information overload, task environment, activity space, problem solving.

Réflexions sur la surcharge cognitive. Cet article aborde le problème de la saturation cognitive tel que les individus le vivent au quotidien sur leur lieu de travail. Il examine d'abord une série d'hypothèses sur les causes de ce phénomène : trop d'information en "push" et en "pull", le multi-tâche et les interruptions, un environnement de travail mal conçu. En passant, il pose une série de questions : la mesure de la qualité de l'information, la forme de sa fonction d'utilité, la pertinence de différentes stratégies de gestion des flux. Ensuite, il cherche à bâtir un cadre d'analyse pour améliorer la conception des environnements de travail. Il enrichit la formalisation de Simon et Newell sur l'espace de la tâche, en montrant que les stratégies cognitives du sujet peuvent l'amener à des détours qui, sans être des tâches proprement dites, sont des reformulations du problème qui facilitent sa résolution. Cette avancée permet de poser plus clairement la

* Dept of Cognitive Science - UCSD
E-mail : kirsh@ucsd.edu

question de la charge cognitive et des coûts cognitifs qui découlent de l'environnement. L'article distingue notamment la charge de calcul, la mémoire, la concentration, le stress. Il montre sur quelques exemples simples comment un réaménagement de l'environnement peut abaisser la structure de coûts pour l'opérateur, et lui procurer une aide dans son travail cognitif.

Mots-clés : saturation cognitive, qualité de l'information, espace de la tâche, résolution de problème, charge de calcul, mémoire, concentration, stress.

INTRODUCTION

Today's workplace is a complex knowledge environment in which the flow of information is mediated by an ill understood array of technologies, at-hand resources, and shifting teams of people. Few of us believe, any longer, that office work is straightforward and procedural. We recognize that people engage in many tasks at once, often in ways that cause interference. People interact with each other and with their tools in little known ways; they constantly develop work-arounds to standard operating procedures, and their primary work space is not confined to the physical region within arm's reach, but is a distributed cluster of 2D and 3D spaces near key resources, computers, telephones and bookcases. Indeed modern workspaces now include virtual spaces – customized computer "desktops" and applications that have their own worlds of organizational structure, information space, and workflow requirements. Given this complexity of tasks and spaces it is no wonder that workers have trouble effectively managing their office activities and coping with information. Email, telephone calls, electronic discussion groups, websites, pushed intranet news, letters and memos, faxes, stick-ems, calendars, pagers, and, of course, physical conversations and meetings, are just a few of the communicative events that bombard today's knowledge worker. The upshot is a workspace of increased complexity, saturated with multi-tasking, interruption, and profound information overload. The effect of this cognitive overload at a social level is tension with colleagues, loss of job satisfaction, and strained personal relationships. (IFTF/Gallup [97] study of Fortune 1000 workers.)

To understand how people handle this bewildering matrix of information and activity spaces typical of modern workspaces requires close attention to the fine grain of interaction. Given the prevalence of multi-tasking and interruption: How do we switch attention from one task to another? How do we maintain control over our multiple inquiries? What do we find intrusive, distracting, or annoying? What are the effects of interruption and what sort of cognitive strategies have people developed to minimize their consequences? There is a large body of psychological literature on attention – both single and dual task attention. But the issues that concern us here, lie as much in the interaction between agent and environment as in the agent's cognitive make-up itself – an area experimental psychologists have spent less time exploring. When people adapt to their environments they not only adapt internally by altering mental processes and behavior, they also change the very environment posing the adaptive challenge. If we are to develop theories of information overload, multi-tasking, distraction, and interruption – all key components of a general theory of cognitive overload – we will have to understand this co-evolution. We will need to understand how people dynamically manage their interaction, how they are cognitively coupled to their environments, and how they structure workflow by using the environment as a cognitive ally.

In this paper I will take a first look at some issues that arise when we set out to design real life environments in which multi-tasking, interruption and cognitive overload are the order of the day. So many specific areas for research are opened by these topics there is space to explore just two:

1. What is cognitive overload?

2. How can an understanding of the cognitive workflow in environments lead us to design better workspaces?

PART ONE: WHAT IS COGNITIVE OVERLOAD?

Many of the consequences of cognitive overload are well described in business studies. In "Dying for Information? — an investigation into the effects of *information* overload in the U.K and World-wide", [Waddington, 96] a 350 page report based on a survey of 1,313 junior, middle and senior managers in the U.K, U.S, Australia, Hong Kong and Singapore – the key findings were:

- two thirds of managers report tension with work colleagues, and loss of job satisfaction because of stress associated with information overload.

- One third of managers suffer from ill health, as a direct consequence of stress associated with information overload. This figure increases to 43% among senior managers.

- Almost two thirds (62%) of managers testify their personal relationships suffer as a direct result of information overload.

o 43% of managers think important decisions are delayed, and the ability to make decisions is affected as a result of having too much information.

o 44% believe the cost of collating information exceeds its value to business.

People feel information anxiety[1] and suffer. But as this study suggests, the focus of overload studies, so far, has been on the consequences of *information* overload. Typically this term has been left sufficiently ambiguous that it is not clear whether it includes other factors such as the increased number of decisions knowledge workers must make, the increased frequency of interruptions they confront, or the increased need for time management in everyday activity – the relentless need to be efficient.

Cognitive overload has something to do with all these concerns. Everyone has so many tasks and obligations that multi-tasking is our way of life. Information is relentlessly pushed at us, and no matter how much we get we feel we need more, and of better quality and focus. Our workplaces are supposed to help us cope with these problems. But our tools and resources remain inadequate. We can turn the ringer off our phones, we can close our doors, we can auto-filter our email, we can personalize search engines, ask people to honor privacy, and so forth. But blocking out sacred time segments or sealing ourselves off from outside contact and even filtering email is not a serious solution in most organizations. And where it is acceptable, it still leaves unaddressed the overload that arises from multi-tasking, interruption and information overload that we create ourselves in having to decide how to manage our desks, files, computers, and different projects. The increase in cognitive overload seems an inevitable consequence of the complexity of our information intense environments.

Let us look at the components of overload more closely. We consider four systems of causes: too much information supply, too much information demand, the need to deal with multi-tasking and interruption, and the inadequate workplace infrastructure to help reduce metacognition.

---

[1] Information anxiety is the overwhelming feeling one gets from having too much information or being unable to find or interpret data. Wurman [90] writes: "Information anxiety is produced by the ever-widening gap between what we understand and what we think we should understand. It is the black hole between data and knowledge, and it happens when information does not tell us what we want or need to know."

1.1 Supply-Related Overload

Following modern conventions we can distinguish two forms of information supply.

- *Pushed* information is information arriving in our workspace over which we have little short term control – the memos, letters, newspapers, email, telephone calls, journals, calendars etc. that land in one of our inboxes.
- *Pulled* or retrievable information is information we can tap into when we want to find an answer to a question or acquire background knowledge on a topic. We have greater control over pulled information in that we intentionally seek it. But it resides in vast repositories such as libraries, online journals, filing cabinets, newspapers, archived discussion groups, our own email and of course the web. At a more interactive level, discussions with colleagues and chat requests in online discussion rooms are additional examples.

Both of these types of information are part of the great supply of information that we must decide whether, how and when to use.

1.11 Oversupply of Pushed Information.

Here is a standard case of pushed oversupply and the activity it spawned. After 10 days of travel a colleague of mine returned to his office. He collected his paper mail from the mail room, went to his office, found on his chair another stack of paper mail far too large to fit in his mailbox, discovered 290 email messages in his inbox, and listened to 14 telephone messages on his answering machine. After scanning the topics of his email, he checked 21 in detail and quickly answered 6, he returned to his telephone messages, answered 4 that were still timely, and then he sat on the carpeted floor, pulled over his garbage can and began tackling his paper mail. As he worked he started building piles on the floor. Journals and magazines for his lab went "there", requests for article reprints he put over "here", he began filling a large manila envelope with letters and receipts he was going to take home. Newly arrived software went beside the journals for the lab, and so on. Halfway through, he stopped to read 30% of a newsletter before trashing it. When he was largely finished, he had 9 untidy piles in different parts of the room – mostly on the floor, and then he left, carrying the pile "designated" for his lab.

No knowledge worker needs reminding that we are bombarded with information. It comes in all categories of urgency, media, size, timeliness, complexity and value. To deal with it we have to make decisions. Is this garbage? Might it be useful? When? Where should I put it? Must I make a new file or new category for this? Can my colleague use this? Making these decisions carries a cost in time, effort and stress. Even if we have a system for making such decisions we still must scrutinize each piece of information, categorize it as of value for this or that project and consider what to do with it. Too often information falls between the cracks of our classifying scheme and we are forced to go through the challenging process of creating new indices and categories, or uncomfortably stretching old ones. Moreover, whenever we create a new category or stretch an old one, there is the danger that where we place the information – on our desk, in our filing cabinet, in a computer folder – will be forgotten the next time we look for it. All this is stressful. But particularly so, the less one has a system for dealing with pushed information and the more one must make *ad hoc* decisions for each incoming piece. The psychological effort of making hard decisions about pushed information is the first cause of cognitive overload.

1.12 An oversupply of Retrievable Information

The information arriving at our doorstep is a fraction of what we need for our work. We constantly consult our files, the Web, our colleagues, libraries, online discussion groups, journals, etc. for more. Another source of cognitive overload stems from the effort of performing effective search. We always seem to want more and better focused information. And no matter what we have found so far, most people harbor a lingering belief that even more relevant information lies outside, somewhere, and if found will save having to duplicate effort.

The problem is bizarre. From an economic viewpoint, if information were like other commodities then beyond an initial threshold we ought to find both the marginal cost and the marginal value of the next piece of information falling with quantity. The more that is produced the less it ought to cost per unit, and the more we accumulate the less additional pieces ought to be worth. Hence our life ought to get easier in information rich environments. The missing premise is quality. In economics it is assumed that it is possible to increase the quantity of a commodity available, in this case information, without reducing its quality. This may be possible in principle, but we all know it is not so in practice. There is evidence that the ratio of high quality to low quality information is falling. See figure 1. The increase in low cost information now readily available with the web has massively outpaced the increase in quality information available. This drives up the individual cost of search for quality information. Where before we turned to trusted information sources, such as refereed journals or quality magazines, we now do more of our information hunting ourselves. Yet no search engine seems to return hits with sufficient precision to save us from having to browse dozens of useless pages in our effort to berry pick the best items. Substantially the same holds true for digital and physical library searches now that these have increased so substantially in size. The result is that we spend more time searching than we feel we should. Time is wasted and jobs that must be done are left undone. More stress. The second source of cognitive overload, then, stems from our effort to cope with the uncertain quality and relevance of our information supply. There is more search and less satisfaction.

In the next section we explore a further consequence of increased information supply: a partially justified belief that there are always higher quality facts somewhere other than where we are looking.

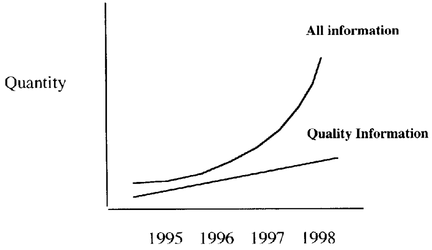


Figure 1. Here we see that in the last few years the amount of information has been rising exponentially while the amount of usable or high quality information has been rising only linearly.

1.2 Demand Side Overloading

Demand side overloading is the consequence of the complexity of our desire function for information. Uncertainty about how much information will be needed, when it will be needed, and how valuable it will be, leads to a complex demand structure that interacts in peculiar ways with information supply. If we treat having the information we think we need at the time we think we need it as a problem in inventory control then we can understand demand side overload as the product of computational complexity. Knowledge inventory control is simply so complex that knowledge workers cannot optimally solve it. Instead they rely on reactive methods of coping rather than careful planning.

Every task a knowledge worker is involved in has its informational requirements. Sometimes these can be satisfied without knowledge retrieval; the worker already knows what must be known, and so he or she can complete the task without searching or soliciting information. More often, though, additional information or knowledge is required. Economic analyses of the value of information have shown that if the cost of missing a piece of information or knowledge is known, it is possible to calculate the time a rational agent ought to spend in securing it[2]. If all the information required by the totality of one's projects is specified, and a value is assigned to each piece, then there is a well defined function determining the optimal allocation of time to be apportioned to any particular information search.

Several factors upset such normative analyses, however. First, knowledge workers rarely know how valuable the information is which they do not have. The theory of information value cannot offer guidance if an agent cannot make a reliable estimate of the value of information. Without a well defined value function, economic analyses are of little use. This occurs whenever a knowledge worker does not know what he or she needs to know to do a task well. For instance, if I wish to write an essay do I need to know the ins and outs of MS Word? It really depends on what I end up doing in that essay. This may be hard to know in advance. How many figures will I have? How many formatting styles? Will I try to automate making my citation list, a table of contents and so on? Normally, we cope with these sorts of problems by using online help, a tool in well designed programs that allows us to get information just at the moment we need it.

Online help is a nice example of a tool for just-in-time learning that does not require going very far outside of one's current environment of activity. It is an excellent example of improved infrastructure. But even with online help there are occasions when it would be unwise to interrupt our task to get the information we need. How deeply into the help system will we have to go? The deeper we go the more we are apt to lose the thread of our composition. In such cases we would like to already have the relevant knowledge.

Moreover, outside of computer environments there is rarely anything analogous to online help. This raises issues associated with the setup cost to starting a search. Where will we find the information? By calling someone? By looking in a manual on our shelf? In someone else's office? Even in cases where this process has

---

[2] For instance, in Minimizing Information Overload: The Ranking of Electronic Messages *Journal of Information Science* 15 (3) 1989, pp. 179–189. Losee defines a "formal economic rule for deciding whether to examine a message: a message should be selected for examination if the cost of doing so is less than the cost of not doing so." p. 183.

been streamlined and there is easy access to information it can still be hard to estimate its value because the value of information is non-monotonic and non-linear. A little can go a long way, or a little knowledge can be a dangerous thing. It all depends on the task. Imagine trying to fix a jammed photocopier. Sometimes you just need to know a little about photocopiers, sometimes you need to know a lot. So even in tasks where we have prior acquaintance with the task it can be hard to know when you have enough information and when you should continue hunting. What you need might be just around the corner. On the other hand, it might be rather far away. In fact, it might not be in the manual or in anyone's head at all! How are you to know? How long should you keep looking before you regard your lack of success to be a sign that the information is not around? It is for such reasons, that coming to a task with a little knowledge already can make all the difference. It is extremely valuable to know the lay of the land in advance. Unfortunately, it is not always obvious when you know the general layout adequately. See figure 2. How nice it would be if we had enough infrastructure to know the complexity of problems and the amount of information required to solve our problems.

I have been arguing, so far, that imperfect information about the value of information makes it hard for knowledge workers to develop a coherent demand function for information, and so to sit down and plan their information gathering strategies. A second facet of their workplace that further complicates developing a clean demand function – and so developing a usable method of knowledge inventory control – is information timing. When does the knowledge worker think that he or she will need the information? Which tasks do they think they will need information for in the next weeks? A knowledge worker may know that he is to deliver a major report in a month, but not yet know whether he will have dozens of different sized reports, memos and discussion papers to deal with in the interim. Or whether he will be reviewing papers, or mostly doing other things. Since most knowledge work requires information gathering it would be most efficient if any information that would be useful for *both* the interim tasks and the major report would be gathered first. In the absence of knowledge of what these tasks are, such a calculation cannot be made and information scheduling becomes ad hoc. It also complicates the filtering function to be used for pushed information. For if there is a chance that incoming documents may soon be useful for new assignments it is rational to keep them around rather than to trash them. Because most documents have a shelf life, and lose their value over time trashing is a key component of information inventory control. But again uncertainty makes it hard to apply a useful trashing strategy.

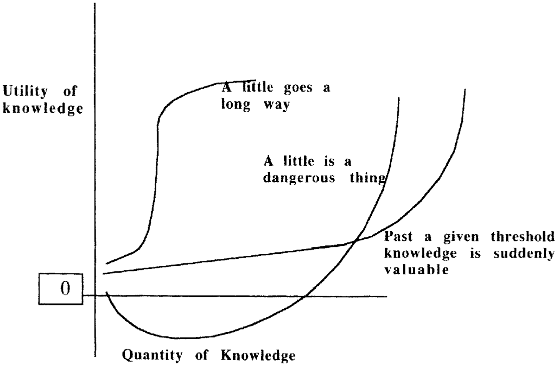


Figure 2. For a given task it is hard to predict the utility structure of information. Here we see three different ways quantity of information affects performance. Utility is here assumed to mean how valuable information is in improving performance. Note that initially information may increase, decrease, or have no effect on performance. Put differently, more or less knowledge may be required for even minimal performance. It all depends on the task and the problem instance.

1.21 Observed Information Inventory Control Strategies

Given the resultant uncertainty in the demand function for information and knowledge, it is not easy to tell knowledge workers the strategies they ought to use for information inventory control. As a descriptive matter though it is common to find people following a few obvious but rather different information gathering and information accumulating strategies. Well designed environments would provide infrastructural mechanisms that would improve performance for each of these different strategies. Here are several of the most prevalent strategies.

Blind Accumulation: If any piece of information might suddenly become critical, it is rational to accumulate all information that might have future value. Naturally, this is rational only if the information can be found later when it is needed; but false beliefs about retrieval abound. The result is overstocking. The cost of overstocking depends on how the information is stored. If it is filed then such people will spend an inordinate amount of time filing. If they save their filing jobs to the end of week or the month, then piles will accumulate in their office.

Just-in-case Learning: Some people like to know enough to be prepared for anything. Just-in-case learning – or just-in-case research – is the result. It is the gathering counterpart to the blind accumulation strategy. In just-in-case learning the emphasis is again on knowledge that might be useful later, particularly if it has a certain probability of having to be acquired sometime. Again, though, because it is impossible to predict exactly when it will be needed there is uncertainty about the current value of information. Conservative knowledge workers may feel the best plan is to always `be prepared'. Hence they will spend an inordinate amount of time reading journals and magazines when they arrive, and doing ancillary research in the library even when they could be extremely focused on a narrow task. Our standard school system is based on this model of learning, coupled with the belief that a certain level of background knowledge must be achieved before any knowledge of a topic can be acquired. E.g. math is necessary for physics or modern economics.

Surface Clutter: Accumulators try to keep all information somewhere, preferably in an easy to access region. But if some of the accumulated information has a shelf life, or if it is expected to be useful in the near future, it will often be "shelved" nearby, so that it may be readily found, may be noticed opportunistically, and the costs of "serious" filing can be saved. As piles accumulate the occasions for creating combinations defining new "overlap" or *ad hoc* categories increase. As this matrix of overlapping categorization deepens it becomes harder to pay the break up cost, leading to an ever more daunting task of filing, if it is ever undertaken at all. Frequent trashing can cull this matrix, but there is also a cost to stirring up the categorization scheme. Understanding how clutter arises and exploring what can be done about it is a rich area for research.

Just-in-Time Information Gathering: A further information strategy it is rational to adopt when you are uncertain about your demand for information is to ignore all information needs except those you know you need for your current task. Rather than collect information on speculation that it might be valuable later, collect information just in time, when you can be precise about what you do need. This is a local maximizing strategy because you guarantee high average value for each piece of information. The danger is that you may not be able to find crucial information in time. Since some information gathering requires advance preparation, you may not know enough to use the information you can find quickly unless you mastered enough background information earlier. Moreover, you risk settling into local optima that are far short of the global optima because you are not maximizing with respect to your overall information needs. In focusing so narrowly you may be myopic.

Trashing strategies: Frequently, knowledge workers temperamentally disposed to just-in-time gathering strategies also are effective trashers: they dispose information soon after it has been used. Just-in-case gatherers are often less effective trashers. It is an empirical topic of some interest to determine the different strategies people have for throwing documents out.

Although all these strategies have been described discursively it is a small matter to restate them formally in terms of standard value of information theory plus individual differences in risk taking disposition.

1.3 Infrastructure Failure

There will never be a completely satisfactory solution to the problem created by an overabundance of information. But we can hope to design operational environments which increase our effectiveness, whatever our strategy, while reducing stress. In addition to information overload, however, there are two other major causes of cognitive overload we must design for: interruption and multi-tasking[3]. They are related.

As a consequence of multi-tasking, agents are constantly interrupted. They begin one task, the telephone rings, they are interrupted when they answer it, and then, if they are not drawn off to a new task, they try returning to the original task. Analogously, they begin the task of writing a report, discover they need a reference, or they realize they need to find out what has already been written on a topic, and off they go in search of new information. This too is an aspect of cognitive overload. A well designed workspace would have tools and resources to minimize the cost of exiting an activity and re-entering it later. It would store much of the worker's task state in some convenient external way. In particular, it would make it easy for the worker to recover his or her place in the task. This means encoding, in some non-overwhelming manner, knowledge of where they had been in the task before the

[3] Multi-tasking is a term drawn from computer science to refer to systems which handle many tasks seemingly at once. An operating system multi-tasks by rapidly switching between tasks and placing in an intermediate store all state knowledge required by each task. It therefore interrupts each task at a stable moment, and then later swaps back the state it was in before the interruption and carries on from where it was. Humans need to be able to stabilize their state knowledge when they are interrupted or they risk losing their place when they attempt to pick up the task later.

interruption, what they had been intending to do next, and what they had learned that was relevant to deciding how to proceed. So far, our environments do little to support these facets of task state. They lack adequate scaffolding. Hence, we are obliged to remember as much state as we can by ourselves. The result is more overload.

The topic of how to design environments to reduce interruption is an area of major theoretical interest beyond our purview here. One important facet of design to discuss now, however, is task collapsing: if we restructure the workspace to integrate several of the different tasks knowledge workers have to perform they needn't be interrupted quite as frequently.

A word is in order. Interruption occurs when we shift from one environment – one task context – to another. Upon exiting we must store that task's state and upon entering another task context, another environment, we must recover its task state. If we can accomplish our tasks in fewer environments we may reduce the number of interruptions we encounter. At a minimum, we may hope to reduce their disruptiveness.

To see clear examples of task integration it is constructive to observe the history of software development. As successful software products evolve they tend to include more and more features that address functional requirements of users. This may often seem like feature bloat – a negative consequence of trying to address the needs of uses with different levels of expertise – but in virtuous cases a program can be expanded into completely new areas so that a task which before required running two programs can now be run within one.

For instance, several graphics programs have now included enough picture manipulation tools to duplicate the functionality of dedicated photo manipulation programs, such as PhotoShop. This simplifies activity because now a user does not have to leave one program with its features and feel, export a file to a new format, open it in a new program with a completely different layout and workspace, and work on it there. It is particularly helpful since we rarely can anticipate all the illustration tasks we wish to perform and all the photo manipulation ones. Typically, we try something, see how it looks, undo it or keep it, and then go on. Our natural workflow is data driven. Whenever we are forced to move between different programs, however, we are obliged to modularize our composition process as much as possible so as to stabilize it. Since this is not, in general, the way we work we find ourselves going back and forth between the two programs as we decide to use a pencil and then an airbrush to touch up our picture. This is disruptive. The effect is more swapping of memory state, more cognitive effort and ultimately more stress than doing the two tasks in a single program.

Enhancing the functionality of environments by absorbing entire activities is potentially a powerful method of decreasing interruptions. A second way, is to collapse multi-step operations into single step ones, so that what before required five steps can now be done in one step. This too is a form of interruption reduction since every time we have to *plan* a lengthy sequence of actions to accomplish a sub-goal we are being interrupted from our main goal by the need to enter a planning phase. Reducing the need for planning sub-goals saves mental effort.

As useful as both these modifications are there comes a design point when there are so many tools available that our environments lose their simplicity and the cost in added complexity outweighs the benefits of convenience. Another form of feature bloat. A case in point can again be found in the history of modern graphics programs, such as PhotoShop, Canvas, Micrographics Designer and Illustrator. With each successive version these packages have added more functionality to the point, a few years ago, where it became difficult for the average user to know where to find all the buttons and tools needed to perform even simple tasks. Obviously a redesign was required.

This redesign occurred soon after. Software engineers struck on the idea of *modes*. In a mode, only certain of the tools present in a program are visible at any moment. Others can be invoked by pressing an icon that causes a whole tray of tools to appear. These tools are context sensitive as well. If we are working on a bit map then standard illustrating tools are not visible; only bit map tools are. But shifting back to drawing does not require saving the file or changing the appearance of the workplace. We just click on the drawing toolkit icon. Admittedly, a neophyte to such programs will still not know where to look to find the tools he or she wants. But for that matter, she will not even know what tools might be available. It is simply not possible to find effective designs for both advanced users and beginners. For that reason our focus here is on intermediate users and beyond. But, we will not go far wrong in assuming that ordinary people are near expert in their daily activities. So to design everyday environments, rather than software environments, we may assume that we are looking for ways of enhancing the activities of everyday experts.

There is much to be learned about the objectives of redesign by studying the details of software evolution. My main point, though, is theoretical. By and large, existing workplaces have failed to keep pace with the incredibly fast rate of change in work requirements.

The shift to knowledge work and the remarkable rise of the World Wide Web has created a new set of informational demands on workers which is leading to ever increased cognitive overload. I believe it is possible to find cognitive principles that can inform design change and reduce cognitive overload. But even in the field of Human Computer Interaction, where such concerns are foremost, our understanding of these principles is still in its infancy.

Many theorists agree that, as a first step, however, we require some account of the cost structure, in cognitive terms, of our workflow. A fundamental obstacle, however, is that the key concepts of workflow, particularly *cognitive workflow*, *task context*, *environment* and *activity space* have yet to receive adequate elaboration. We can all agree that the objectives of designers is to realign the cost-structure of the environment, thereby reducing the stress and cognitive costs of performing a job. But without a theory of what occurs when this happens it is not possible to let theory drive design.

My goal up until now, accordingly, has been to show just how complex the informational world is in which we live, and to introduce the argument that to make serious progress in the design of better workspaces we need a new theoretical understanding of our activity space and our dynamic relation to our environments. The line I shall take next is that to design for cognitive overload we must understand the full set of cognitive relations we have with our environments. An important step is to clarify what an environment of activity is and the diversity of ways we connect cognitively with such environments. It is to that I now turn.

PART TWO: COGNITIVE WORKFLOW AND ENVIRONMENTS

As can be appreciated from our brief discussion of software adaptation, the goal of good software designers is to create work environments which complement and simplify users' workflow. The same applies to designers of everyday work environments. But what does that phrase "*complement and simplify users' workflow*" really mean?

One thing it means is that a well designed piece of software should provide a good work "environment" – an integrated workplace – for users to perform all the tasks and sub-tasks that go into performing a job. To a first order, workflow is the way users move through these tasks and sub-tasks. For instance, if their job is to write an essay then the major workflow tasks involve outlining possible topics, sketching ideas, rearranging text, copying and editing, erasing, finding synonyms, formatting, adding citations, placing pictures, spell checking and the like. A good word processor ought to create a work environment with the appropriate tools and resources to facilitate these tasks. Because a digital medium makes it easier to cut, paste, copy, store, tag, and manipulate, a word processor ought to provide a better work environment in which to work than paper and pencil. The workflow of composition should be simplified and complemented in digital media. Hence digital media such as word processors constitute a better environment for the activity of writing essays.

As word processors mature, however, designers are beginning to appreciate that there is more to the workflow of composition than is found in the organization of the sub-tasks which modern word processors support. Word processing is itself a system of tasks woven into a larger tapestry of activities that make up the more complete process of composition. Many of these activities take place outside the computer, on scraps of paper, in dictation devices, in discussion, in front of whiteboards, in using library search engines, in skimming abstracts, in reading and annotating articles and books, in making notes, and even in conversations taking place through email. The workflow of composition is multifaceted. It takes place in many environments, involves many tools, and requires considerable mental effort[4].

Workflow analyses take a more realistic and a more cognitive turn when they expose the fine details about how people exploit their environments to make the cognitive aspect of their activities easier. All too often workflow analyses have been confined to studies of how workers move from one aspect of work to another as they complete a task. As we see in composition this may involve acting in a single environment, in many environments, or in making linkages across environments. But a more complete analysis of workflow will not only identify the principle components of a task – the sub-tasks and activities that constitute it and the different environments in which they occur – it will also track the lines of information – both data lines and control lines – that link the component tasks and permit the author to keep track of her goals across environments. What representations do people use when they compose? What

---

[4] It is not easy to predict the details of how word processors will change to accommodate this more complete reality of the composition workflow. But we can be confident that current word processors are just a moment in the design evolution of composition tools. The considerations that are driving them to evolve have to do with understanding what people do when they compose, and how they move from one aspect of the composition process to another. For this reason it is a good bet that tomorrow's word processors will have to provide better facilities for taking notes on digital library materials, and better facilities for displaying these notes and reference materials so that writers can cut and paste between them.

representations do they use to keep track of where they are? What activities do they perform to bring different media together? How do they coordinate their activities? These various activities of planning next moves, clearing off clutter in their environments to get on with their sub-tasks, marking intentions to make it easy to pick up where they left off, tracking what they are doing, highlighting, annotating, talking to themselves as they work – all these and more, constitute elements we must study if we are to begin to paint a more complete account of the cognitive workflow occurring when people compose documents.

2.1 Environment as Activity Space

I have been suggesting that if we closely track the workflow of individuals we see them operating in several different environments, attempting to coordinate the key processes in each by a variety of techniques ranging from talking to themselves, to annotating, to making notes, to leaving cues and markers and so on. More often than not people do not complete one task before initiating another, and they move from job to job, environment to environment, as a function of plans, interruptions, exigencies and habits. The complexity of this trajectory and the fact that on certain analyses agents change task environments without changing physical environments makes understanding workflow at a fine level of detail a daunting job.

Lurking at the bottom of this more cognitive way of thinking about workflow is the concept of an environment as the space in which work takes place. Since the primary goal of designers is to improve the structure of an environment to make it easier to work in, we need to look more deeply into the relation of workflow and environment.

We shall think of an environment as an activity space – originally a physical space but now virtual spaces qualify as activity spaces as well – populated with resources, tools and constraints in which an agent operates. The reason the same physical space can support multiple environments or activity spaces is that the way an agent projects meaning onto a space partly *constitutes* it. A snapshot of a checkerboard along with its pieces can (with some imagination) be seen as a moment in either a game of checkers or a game of chess. How things unfold soon gives the lie to one of these interpretations, but the simple fact is that there is more to activity spaces than physical structure. Agents project meaningful structure onto environments.

This process of constitution depends on having:

1. a task to perform and thus seeing the affordances of the environment through task-colored glasses,

2. a history of experience with similar objects and resources, and thus having a body of associations that enrich the meaning and understood possibilities of each situation, and

3. acknowledging certain norms of work conduct.

Change any of these and you change what the agent thinks he or she may do in an environment, even what they think is possible in that environment, and what they see the consequences of actions to be. In brief, change one of these and you change the activity space – the environment of action.

At the same time that an environment represents a space of possibility it also represents a set of constraints. An environment is the space in which structures are created and actions have consequences. It is the substrate in which new structural and meaningful configurations (situations) can be created and the substrate which constrains the possibilities of creation. Not anything can be created. The work environment therefore constrains both what it is possible or acceptable to do, and what happens as a result of performing actions. It is partly the product of an agent's projections and partly the product of underlying causal realities.

It is important to understand the parameters of an environment since any effort at designing environments to minimize interruption, disruptiveness and to facilitate information management will focus on manipulating these design parameters. A natural way to capture the sense of constraint and possibility inherent in an environment is to enumerate the set of actions it supports. What can an agent do while there? And what happens as a result of those actions. As noted by Newell and Simon, [Newell 82,90, Newell and Simon 72, Simon 73] this characterization, as it stands, is too free; it lacks focus. We don't want to include in an environment's activity space *all* the actions an agent can perform there. In physical environments, for example, agents can move their arms, wiggle their noses, jump up and down, push objects at different speeds, as well as numberless other actions. Sometimes their actions result in a change in themselves (scratching makes me less itchy) or in a momentary change in their relation to their environment (running on the spot, walking to a new place). At other times, the focus of action is on the objects in the environment, and change consists in having these objects occupy new positions or orientations (move the frying pan) or occupy new states (crack the shell, egg is fried). We need to narrow our notion of the activity space to the space of actions that are relevant to the task at hand. Actions that make sense to a purposeful agent.

Newell and Simon restricted their notion of environment to the task environment. They included in a task environment just those actions which might make a difference to the task. This worked quite well for formal problems such as Tower of Hanoi, where the environment was impoverished. In these stylized environments the actions deemed essential to the task environment all have to have a *measurable effect* on performance – on the time, energy, or number of actions an agent would need to complete the task. They *cannot be too general or task independent* for then they would have nothing specifically to do with the task. For example, breathing, blinking and perspiring are behaviors which when not undertaken may drastically affect performance. But their contribution is more to the general condition of the agent rather than to anything task specific. They are background actions which contribute to the creature's being in its normal state of being, hence in a task ready state, but are not themselves part of any task in particular. At the same time, actions *cannot be too idiosyncratic*; they cannot be so specific to one agent, that any other agent also performing similar actions in similar task circumstances would neither improve nor worsen its position in the state space. Performing a chant before hitting a tennis ball may be "necessary" for a certain player, but it is certainly not part of the activity space of tennis[5].

By focussing on task-relevant actions, Newell and Simon attempted to operationalise the idea that one's environment of action – one's task environment – is somehow tied to projected meanings as well as to what potentially works in that environment. The activity space that matters when performing a task is the one which is somehow connected to accomplishing the task in an allowable way. I believe this is substantially correct.

But Newell and Simon sacrificed too much for their formalism. They adopted a narrow view of both the creative possibility of humans and what it means for an action to be connected to a task. We know that people constantly exploit the affordances of objects to do things they believe may advance their case even though these behaviors often cannot be found in the canonical action repertoire. Similarly, we know that people perform a host of actions that make no sense from a pragmatic as opposed to an epistemic viewpoint. Not all actions can be seen as potentially bringing one physically closer to the goal state. Rather they help bring one epistemically closer. They serve a host of cognitive functions that are not merely

[5] This tripartite description of the conditions determining the actions to be included in the action repertoire linked to a task environment is drawn from Kirsh, D., "Adapting the environment instead of oneself", *Adaptive Behavior*, vol. 4:3/4, pp. 415-452, 1996.

idiosyncratic; they are cognitively useful given particular strategies, ways of looking at problems, and cognitive styles. So although these actions again fall outside the canonical action repertoire they are connected to accomplishing the task.

It is this last collection of actions that makes analysis of environments especially difficult. In the Tower of Hanoi, for instance, all the rings are supposed to be resting on the three pegs at every moment. There are no extra degrees of freedom in where they can go and what you can do with them (unless you count holding them in your hand before placing them on a peg as something you can do). The rules of the game even outlaw using pencil and paper to help you decide. After all, the Tower of Hanoi is a type of memory task. But no one prohibits muttering to oneself. Such muttering can potentially improve performance, since it engages the phonological loop and so can offer a few bits of extra memory to encode useful task information. Effective encoding in this loop can reliably improve performance. So it ought to be part of the activity space. But by Newell and Simon's narrow definition it is not an event in the task environment. Since, virtually every task in the natural world offers agents the possibility of using local resources to cleverly encode task useful information; it may be difficult to keep a tidy notion of task environment while viewing each such resource use as a task relevant event. But it is hard to see how we can deny that encoding helpful knowledge in an environment is part of task relevant performance and so an element of the task environment. All potential actions that can reliably improve performance are by definition part of the task environment.

Analyses of task structure are further complicated because in most tasks people undertake outside the laboratory, the rules of engagement are not well defined. The actions and meanings that are projected onto the environment do not flow from the agent or the task setup unconstrained by the history of agent environment interactions. They are somehow "negotiated". I habitually use a slipper to stop our screen door slamming in the wind, although that is not its orthodox function. In cooking dinner, I find it convenient to use the large wrapping paper my fish comes in as an interim plate to hold the fish once it is floured. Not only does this save me from washing an extra plate, it gives me a surface that is large enough to lay out all the fish so that they when I flour them they do not touch each other. If I were to pile them on a plate they would stick together. The moral is that people co-opt the function of everyday items to help them with their tasks. Sometimes their creativity saves them physical resources (e.g. re-using wrapping paper and so saving a plate), sometimes it saves them physical effort (e.g. not having to open the door each time it slams shut), sometimes it saves memory (e.g. not having to remember the task state by encoding information in the layout), sometimes it save mental effort or computation (e.g. preparing an assembly task by laying out the pieces to be assembled in a geometric pattern that shows one where to put the next piece.) An adequate definition of environment or task environment should allow any of these creative uses of tools and resources to be part of task performance. It should recognize that people make cognitive use of their environments whenever they act.

These ideas about the activity space of tasks has clear consequences for design. Well designed environments ought to take note of the cognitive needs people have in performing their tasks and build in *scaffolding* that simplifies the way people make cognitive use of their environments to increase task reliability. There is no end to the variety of methods that can help, so this is not a closed design task, where there is an optimal solution. Nonetheless, we can start with a description of the cognitive or epistemic actions afforded by an environment and work from there to find ways of overcoming some of the physical and cognitive constraints of the given environment.

2.2 Cognitive Workflow

When we observe the fine grain of activity occurring as a person performs a task it is evident that much of what they are doing is not directly concerned with improving their "pragmatic" position. [Kirsh forthcoming, 98, 97, 95, Kirsh & Maglio 95] Some of their actions serve an epistemic function, and some serve a complementary function, helping agents to coordinate their internal and external activity[6].

A particularly clear case of an epistemic function is found in the computer game Tetris. Good players will rotate a piece very soon after it appears at the very top of the board if by so doing they may more easily discover what sort of piece it is. Their action unearths information. Similarly Tetris players will rotate pieces externally when they are trying to decide what to do next rather than rotate the mental image of those pieces. Besides saving the mental effort of image rotation, physical rotation is almost three times as fast. Players are able to more quickly enter the mental state of knowing what their piece looks like when rotated.

---

[6] Several nice examples of complementary actions are found in the way we use our hands to help us think and see. To count closely packed items, especially if they are identical in appearance, we often need to point because our visual systems easily skip items. To compare items we like to pick them up and put them beside each other.

A particularly clear case of complementary actions that help coordination is found in young children when they count objects in story books. Unless they coordinate the rhythm between the way they point at objects and the way they count out loud their error rate soars. Internal counting (which is coupled to external counting) must be coordinated with external pointing (which is coupled to visual tracking). Disrupt this rhythm and they lose their stride, and falter.

A more sensitive account of workflow ought to describe the cognitive and interactive processes that are occurring when agents perform their sub-tasks and move from one aspect of work to another. I call this type of workflow analysis "cognitive workflow". It is the next step in understanding designers must take if they wish to design scaffolding and other environmental resources that can reduce the cognitive complexity of tasks. It implies that we must understand the cognitive environment the agent operates in.

Many of the activities that are meaningful in the cognitive environment are easy to understand. Returning to the task of composing an essay, we see that agents perform all sorts of meaningful activities that help them process information and help them encode idea fragments. For instance, agents reorganize their references and notes. If their references are physical articles and books, they put them on their desk or on their shelves in one or more piles. As they read them they shift their spatial position, possibly keeping them open at salient places by putting them face down, or by folding or bookmarking them, perhaps interleaving pages of two or more articles to help define an ad hoc category, or possible theme. If the references are digital then they shift them around their computer screen, hiding and occluding some, putting others in new directories.

But many of the important interactive activities we perform in a task are less easy to understand. For instance, in managing the resources available to us, we may develop routines that are idiosyncratic. I mentioned several different knowledge inventory control methods above, but there may be many more such methods we use for dealing with clutter, or helping us to manage our time. All these are additional relations we bear to our environment that we would like support for. And this is before we consider the complications added by working on tasks with others, or the complications added by working with complex artifacts and tools. Since most of our environments include both of these highly complex systems we are constantly relying on a further collection of strategies for coordinating our actions.

To return to composition, in today's world, and probably for some time to come, many of the resources we use are both tangible and virtual. Most books currently reside outside the computer, and we value the affordances this provides, and certainly our colleagues reside outside our computers and most of discussions with them take place non-virtually. But use of tangible resources including people also carries a transaction cost for the modern writer. For now we must transfer the contents of those books, and conversations, as well as the contents of any notes and annotations we wrote down, from outside to inside our computer. No doubt an effective way to meet this challenge will be worked on for several years. It is not a simple problem to be solved by digital scanners or video cameras. For our goal is to capture the content arising in the activity of composition as it is distributed over many spaces and many environments . Each of these environments – virtual and physical – has its own functionality and each has its own special resources that we have learned to use. Accordingly, the first step in redesigning the composition landscape is to create a map of the cognitive workflow. This involves tracking the trajectory of these highly situated meanings through these multiple environments. It shows how work is coordinated over many functional environments and how agents manage the resources available in each environment to maintain control of their overall task.

2.3 Altering the Cost Structure of an Activity Space

The principle objective of redesigning environments is to improve workflow – both cognitive and physical. Improved workflow ought to lead to reduced cognitive overload, since one key component of a well designed environment is that it minimizes the cognitive effort agents must exert when performing their task. To make this idea of improving workflow and minimizing cognitive effort more operational, however, requires quantifying the cost structure of an environment. It is here that the idea of treating an environment as an activity space proves useful. Given the tools and resources available in an environment we can try to estimate how easy or difficult it is to perform an action. Once we understand how to improve the cost structure of particular environments – of how to lower the cost of managing activity in that environment – we may then ask how to improve the design of the overall workplace, since our actual workplace is really a superposition of many specific environments which we slip between. Put in a slightly different way, there are many distinct interpretative frameworks we operate with during the course of a day or during the course of an activity, and these frameworks interact with our workspace to create a theoretical construct we are calling an activity space. As designers we must analyze activity spaces separately before trying to come up with an effective design for their superposition.

Returning now to the question of operationalizing the cost structure of a specific activity space we can ask of a specific environment how hard is it to perform a task – for instance, to write a note in it. If we assume a normal physical environment, then presumably as long as there is pen and paper around and there is a flat surface to write on, it is easy to write a note. Of course there is more to it than this, and calling a task easy is a purely qualitative judgment. But at least we differentiate note-taking in minimally equipped environments from note-taking where there are no convenient affordances for writing – no flat, clean surfaces, no places to rest. Without the necessary affordances the cost of writing in E_1 is much higher. Not high, perhaps, for making simple marks on an existing note, but for any extensive amount of writing it may be extremely difficult or painful to produce legible output. The cost structure of E_1, for writing, then, is non-linear. To write large notes is far more difficult than writing the same number of words in several small notes. Add a table and the overall cost structure is lowered and flattened. Of course, so far we have said nothing of psychological factors. What is the content or topic of the note? Jotting down a phone message or writing an essay? Depending on the task demands different factors increase in importance. How noisy is it in E_1? Is writing a note the only thing agents must do at the time? What is the mean time before the next interruption? How demanding will that interruption be?

Or consider the effects of clutter in the environment. If our goal is to take a note in response to a telephone message, how hard is it to find the message pad in E_1 or the earlier note to be annotated? It depends on who is looking. Rooms with stacks of files everywhere are likely to make the task of finding things hard for anyone who has not been involved in creating the clutter. Less so for those who created the mess. The cost structure of retrieval in E_1 depends on what has to be found and where it is located. Even more important, the cost structure also depends on what the agent knows. If the author of a mess were to share facts about his 'indexing' or filing system the overall cost structure of retrieval (by an arbitrary person) might be significantly lowered.

The idea of activity spaces having a cost structure – a metric which assigns particular costs to particular actions – is intuitively attractive but requires clarification and operationalization before it can be put to constructive use. A first step can be taken by distinguishing behavioral measures of environmental goodness from cognitive measures.

On the behavioral side the key questions naturally concern performance related measures, such as time to complete the task, number of errors made at a given speed (speed-accuracy trade-off), the probability of making an error on being interrupted and the cost of that error, the hardest version of the task that can be performed in that environment. See Figure 3. Certainly there may be complex non linear relations between these parameters, but in the simplest case a given environment E_1 is better than environment E_2 if all parameters are better in E_1 and none are worse (the Pareto sense).

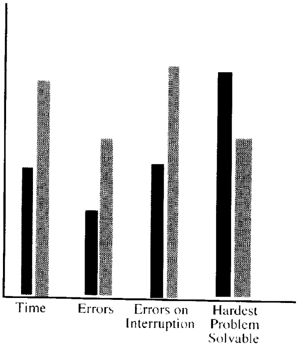
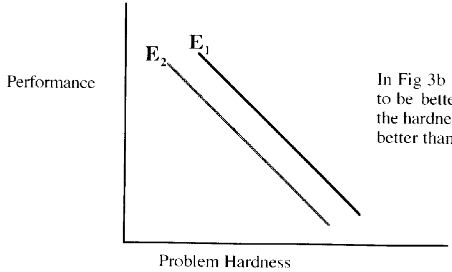


Figure 3. Environments or activity spaces differ in several performance related dimensions.

In Fig 3a we see that the same task can be solved faster in E_1 than in E_2. that fewer errors are made, that E_1 is more robust to interruption, and that harder problems can be solved. The speed accuracy curves of E_1 therefore are better than in E_2.



In Fig 3b we see that E_1 continues to be better than E_2 as we increase the hardness of problems. So E_1 is better than E_2 in a Pareto sense.

Behavioral measures are important for *measuring* cost components. But in fact the underlying factors most directly connected to cognitive overload and to cognitive workflow have to do with cognition. Obviously the factor most directly determining overload is cognitive load. Environments that are better designed than others should allow the same task or task instance to be solved or performed with less cognitive load, measured in terms of computation, memory, amount of sustained concentration, and stress. See figure 4.

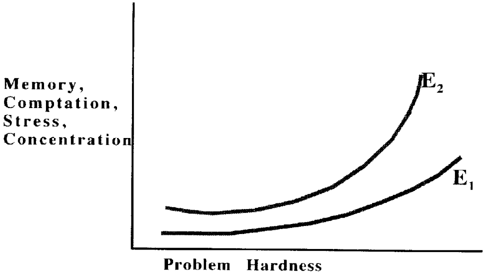


Figure 4. Here we see represented the key cognitive dimensions that determine the cost structure of an activity space. For a given environmental setup we would like to determine the cost in terms of memory, computation, stress and attention or concentration required to perform tasks of varying hardness.

We have all had the experience that a task may be easier to solve if the right artifacts are present. It is easier, for instance, to sum a long column of numbers with a calculator than in the head. It is more reliable to sum a list if the calculator displays the last few numbers just added, since then if you are interrupted you can merely look at the number trail to see where you left off. Similarly, it is easier to sum a long list of numbers if you use ruled paper with small lined columns, since then you will find it easier to keep your numbers more precisely lined up; and that summing numbers that are nicely aligned in turn makes it easier to focus on the appropriate next number to sum, and so to reduce the probability of errors. Over the years innovation and experience has led to the introduction of graph paper with small squares to guide printing. See Figure 5.

Cultures (both national and micro-cultures such as business organizations) accumulate helpful tricks for reducing the complexity of tasks and increasing the capabilities of their members to undertake harder tasks. Sometimes the way culture increase capacity is by introducing helpful representations, or helpful designs. This is a fundamental component of activity space since so much of our activity involves using representations and representational formalisms.

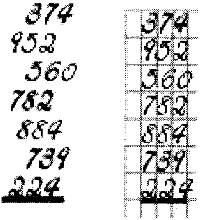


Figure 5. It is easier to sum numbers when they are lined up. Graph paper is a cultural adaptation to help people align their numbers.

It is well known that if the way information is represented in the environment is changed it is possible to change the cost structure of information because some information will be easy to extract or notice – on the surface – and other information harder to extract; that is, more computation will be required to extract it[7].

For instance, there are times when a table of values is a good way to represent the relation that holds between the different cells in it, and times when it is a weak or bad way. It all depends on the information that is to be extracted. A simple example is seen in figure 6. Here the goal is to identify the maximum value. As is clear from the figure, the perceptual action of extracting the global maximum is much easier in graphical form than in matrix form. Global maxima tend to pop out in graphical forms because there are powerful parallel visual processes for noting certain key spatial relations in the visual field. This is especially true for maps where a great deal of Cartesian distance information can be neatly encapsulated in a 2D map. On the other hand, if one's goal is to determine the distance of key stops along the way, a matrix representation which explicitly states the distance in miles between those stops will be easier to work with. See figure 7. Each representation, graphic type and matrix, has its strength and weaknesses. Each makes some pieces of information easy to pick up and other pieces hard.

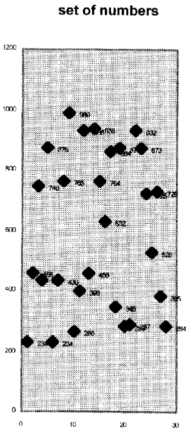
The fact that there are different ways of representing the same information, each of which imposes different demands on the cognitive faculties of users becomes particularly interesting in the course of everyday activity as soon as we ask:

---

[7] For an extended discussion of the differences between information that is one the surface – explicit – and information that is buried – implicit – see "When is information explicitly represented?" [Kirsh 90].

○ how people structure the space around them to encode information which they expect to need soon;

○ what representational artifacts and tools they have to help keep track of useful information.



234
458
746
436
876
234
436
765
989
265
398
936
456
938
764
632
864
345
876
283
287
932
873
725
528
728
385
284

As an example of using space to encode information consider the simple problem of sorting playing cards while you play a game such as gin rummy. Cards that are well arranged for gin rummy display their owner's key sub-goals and plans. For instance, in Figure 8b, it is easy to see that the player has almost won and is just looking for two of the following: a 4 spades, a 3 or 6 of clubs, K of hearts or K of diamonds. In Figure 8b, by contrast, it is much harder to tell the player's goal structure and therefore harder to note what cards are being sought after.

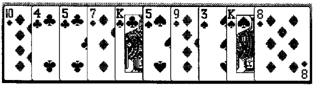


Figure 8a.

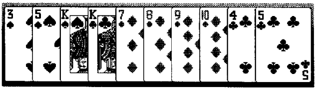


Figure 8b.

Representation is one key method agents have of simplifying their work activity. Another is by learning new practices or algorithms that make it easier to solve problems. For instance, in gin rummy a useful practice is to first sort cards in ascending order regardless of suit, then sort them by suit in ascending order. This practice has the effect of showing off all potential sub-goals, since it lowers that chance of noticing cards of the same value. After sub goals have been chosen agents can re-order their cards as they wish to highlight the plays they are looking for.

In a more serious vein managers in corporations are often taught how to conduct better meetings, how to coordinate the activities of team members so that more is done with less effort and stress. These methods often become part of the corporate culture and new employees are sent to training class to learn them. Typically practices involve use of particular representations or resources. For instance, day planners are widely used to help structure the workflow of their users, helping them to organize their activity throughout the day and record commitments, obligations and comments in a single easy to find place, thereby reducing the stress of having to remember certain facts and commitments and offload both memory and effort. New resources lead to new practices, usually representational practices, and the net result is that the cost structure of the activity space is altered.

To sum up then, several key facts emerge in understanding the cost structure of an activity space.

1. the relevant activity space must be defined relative to a task or family of tasks

2. there is no absolute measure of the cost of different actions or activities to be had in abstraction of the different algorithms which agents have. Change the algorithm and the cost structure of the activity space also may change.

3. the cost structure is also relative to the underlying skills and cognitive capacities of an agent. People with large working memories or who are expert in a task may experience a different cost structure than novices or people with smaller working memories.

4. The resources available in an environment affect the cost structure only if users have developed means for using them effectively in context. Often this is the consequence of cultural learning. But some of these resources can be analyzed themselves in terms of their own cost structure.

To do justice to the variety of ways individuals and especially groups have of reshaping the cost structure of their activity space, and to understand how activity spaces are superimposed one on the other, is a topic of great interest. One aspect I have not discussed is how agents coordinate among themselves, how they create new types of coordinative structures and practices that simplify tasks and distribute load among individuals and their environment in ways that reduce stress and improve performance. To my knowledge no one has yet discovered how to quantify the impact on cost structure such coordinative mechanisms have. My objective here has merely been to introduce the notion and suggest how it might figure in analyses of activity and cognitive load.

CONCLUSION

Cognitive overload is a brute fact of modern life. It is not going to disappear. In almost every facet of our work life, and in more and more of our domestic life, the jobs we need to do and the activity spaces we have in which to perform those jobs are ecologies saturated with overload. As technology increases the omnipresence of information, both of the pushed and pulled sort, the consequence for the workplace, so far, is that we are more overwhelmed. There is little reason to suppose this trend to change.

On the positive side, though, I have been arguing that designers and participants both have the capacity to reshape work spaces to alter the cost structure of the activity that takes place in those spaces. These efforts to reshape activity space come in three forms.

- Change the physical layout,
- change the methods, algorithms and practices agents use to perform their tasks, and

o change the resources, particularly the resources associated with cognitive scaffolding available in the environment.

By changing the physical layout useful affordances of both a physical and cognitive nature can be brought closer to where agents need them and at the time they need them. The simple fact that these affordances are available at the right moment can help agents notice possibilities they might otherwise overlook. A trivial example is leaving an easy to use calculator by the desk of someone doing their taxes, or by substituting a thin computer screen for the large monitor normally on one's desk. The calculator soon gets put to good use, and the extra space on one's desk soon becomes filled up with helpful paper memos, paper products, books etc that agents find useful for their work, but which before had to be stored or placed off to the side. Similarly, by relocating people with complementary knowledge in easy ear or eyeshot from each other, the result is usually that they consult and help each other more than before. They may also interrupt each other too, so the resulting social and work ecology is not unambiguously for the better for everyone. But the received wisdom of our day is that teams of cooperating agents distribute cognitive effort and thereby reduce individual stress.

By changing the methods, algorithms and practices of agents cognitive load can be reduced because better methods of solving problems -- methods that allow practitioners to solve problems harder, faster, or more accurately -- the chief factors we think are the driving factors of the cost structure of their personal activity space. Examples of such improvements are easy to find. For instance, better techniques for conducting meetings, for personal time management, for recording results, for accessing corporate memory, for dealing with interruptions, and for coordinating activity both at an individual and group level can reduce cognitive load. Often these will require that agents use artifacts they never used before. Day planners for time management, wall charts or white boards for meeting management, new search software for indexing and retrieving knowledge and information (i.e. knowledge management), or learning to use the hold button on the telephone, are all examples of artifacts which when effectively combined with practices can improve efficiency and reduce overload and stress.

These last set of artifacts are examples of the sort of resources often called cognitive scaffolding that designers are eager to create. Other examples are new representational formalisms such as Pert and Gannt charts, better information visualization, better search engines, new devices for non-intrusively capturing and recording views of what was said and done, and even contextualized help in our physical workplace[8]. There is no magic design principle that yields new scaffolding. I have been arguing, however, that at a theoretical level it may be possible to analyze in a more rigorous manner the cost structure of activity spaces and the impact which alterations might have on that cost structure. It is not clear whether these analyses will prove useful to designers before they innovate helpful change or only after their most creative phase and they are attempting to analyze why their idea works and how it might be generalized. But as cognitive scientists we need such analytic frameworks to contextualize our discussion of physical and cognitive behavior. It has been my goal in this paper to advance this discussion further by introducing a set of distinctions and concepts that are consistent with other thinking in the cognitive and biological sciences.

ACKNOWLEDGEMENTS:

The ideas in this paper owe much to helpful conversations with Saadi Lahlou, Aaron Cicourel, Ed Hutchins and Jim Hollan.

References

Berghel, H. (1997) Cyberspace 2000 dealing with information overload. *Commun. ACM* 40, 2 (Feb. 1997), pp. 19-24

Chase, W.G. & Simon, H.A. (1973) The mind's eye in chess. In W.G. Chase (Ed.), *Visual information processing* (Chap. 5). New York, NY: Academic Press.

Institute of the Future & Gallup Organization. Managing Corporate Commiuniation in the Information Age. 1997. http://www.iftf.org.

Kirsh, D. (1998) Adaptable Rooms and Cognitive Workflow, in Streitz, N, et al. (Eds.) Cooperative Buildings. Integrating Information, Organization, and Architecture. Lecture Notes in Computer Science. Springer: Heidelberg.

Kirsh, D. (1997) Interactivity and MultiMedia Interfaces. *Instructional Sciences*.

Kirsh, D. Are all Actions Exploratory or Performatory? *Ecological Psychology*. (forthcoming).

Kirsh, D. (1996) Adapting the Environment instead of Oneself. *Adaptive Behavior*.

Kirsh, D. (1995) The Intelligent Use of Space. *Artificial Intelligence*.

Kirsh, D. and Maglio, P. (1995) On Distinguishing Epistemic from Pragmatic Actions. *Cognitive Science*.

---

[8] So far, we have not begun treating our work spaces as domains that can be enriched with context sensitive help. We rely on co-workers to help us, and occasionally we consult manuals, but to date, we have not harnessed the possibilities of computers to give us contextual help outside of the computers themselves. As our walls become information visualizers we can expect this to change.

Kirsh, D. (1990) When is Information Explicitly Represented? *The Vancouver Studies in Cognitive Science*.

Losee (1989) Minimizing Information Overload: The Ranking of Electronic Messages *Journal of Information Science* 15 (3), pp. 179-189.

Nelson, Mark R. (1994) We have the information you want, but getting it will cost you!: held hostage by information overload. *Crossroads* 1, 1 (Sep. 1994), pp. 16-21

Newell A. (1982) The Knowledge Level. *Artificial Intelligence 18*(1) pp. 87-127.

Newell, A. (1990) *Unified theories of cognition*. Cambridge, Mass.: Harvard University Press.

Newell, A., & Simon, H.A. (1972) *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Simon, H.A. (1973) The structure of ill-structured problems. Artificial Intelligence, 4, pp. 181-202.

Waddington, P. (1996) Dying for information: an investigation of information overload in the UK and world-wide. London: Reuters Business Information.

Wurman, R.S. (1990) *Information Anxiety : What to Do When Information Doesn't Tell You What You Need to Know*. Bantam.