



ELSEVIER

Cognitive Science xxx (2004) xxx–xxx

COGNITIVE  
SCIENCE<http://www.elsevier.com/locate/cogsci>

## Soft constraints in interactive behavior: the case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head<sup>☆,☆☆</sup>

Wayne D. Gray<sup>a,\*</sup>, Wai-Tat Fu<sup>b</sup><sup>a</sup>*Department of Cognitive Science, Rensselaer Polytechnic Institute, Carnegie Building (Room 108), 110 8th Street, Troy, NY 12180, USA*<sup>b</sup>*Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA*

Received 15 October 2002; received in revised form 15 September 2003; accepted 15 December 2003

### Abstract

Constraints and dependencies among the elements of embodied cognition form patterns or microstrategies of interactive behavior. Hard constraints determine which microstrategies are possible. Soft constraints determine which of the possible microstrategies are most likely to be selected. When selection is non-deliberate or automatic the least effort microstrategy is chosen. In calculating the effort required to execute a microstrategy each of the three types of operations, memory retrieval, perception, and action, are given equal weight; that is, perceptual-motor activity does not have a privileged status with respect to memory. Soft constraints can work contrary to the designer's intentions by making the access of perfect knowledge in-the-world more effortful than the access of imperfect knowledge in-the-head. These implications of soft constraints are tested in two experiments. In experiment 1 we varied the perceptual-motor effort of accessing knowledge in-the-world as well as the effort of retrieving items from memory. In experiment 2 we replicated one of the experiment 1 conditions to collect eye movement data. The results suggest that milliseconds matter. Soft constraints lead to a reliance on knowledge in-the-head even when the absolute difference in perceptual-motor versus memory retrieval effort is small, and even when relying on memory leads to a higher error rate and lower performance. We discuss the implications of

<sup>☆</sup> Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.cogsci.2003.12.001](https://doi.org/10.1016/j.cogsci.2003.12.001)

<sup>☆☆</sup> An earlier, much simpler version of this report was presented as an eight-page conference paper at CHI2001. That paper is archived as Gray and Fu (2001).

\* Corresponding author.

*E-mail address:* [grayw@rpi.edu](mailto:grayw@rpi.edu) (W.D. Gray).

# Report Documentation Page

Form Approved  
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE <b>DEC 2003</b>		2. REPORT TYPE		3. DATES COVERED <b>00-00-2003 to 00-00-2003</b>	
4. TITLE AND SUBTITLE <b>Soft constraints in interactive behavior: the case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head</b>				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Carnegie Mellon University, Department of Psychology, Pittsburgh, PA, 15232</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited</b>					
13. SUPPLEMENTARY NOTES <b>The original document contains color images.</b>					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES <b>24</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

soft constraints for routine interactive behavior, accounts of embodied cognition, and tool and interface design.

© 2004 Cognitive Science Society, Inc. All rights reserved.

*Keywords:* Interactive behavior; Rational analysis; Knowledge in-the-world; Human–computer interaction; Embodied cognition; Embodiment level; Soft constraints; Bounded rationality; Interface design; Information access; Eye tracking

---

## 1. Introduction

Knowledge can be in-the-world or in-the-head (see, e.g., Larkin & Simon, 1987; Norman, 1989). A well-designed interface can place knowledge in-the-world so that it is available in a known location when a user needs it. Support for acquiring knowledge from in-the-world rather than retrieving it from in-the-head is widely touted as one of the main advantages of direct manipulation interfaces over command language ones (Frohlich, 1997; Hutchins, Hollan, & Norman, 1985; Shneiderman, 1982). This cognitive engineering view is congruent with accounts of embodied cognition that suggest a privileged status for perceptual-motor effort compared to retrieval effort (Ballard, Hayhoe, & Pelz, 1995; Ballard, Hayhoe, Pook, & Rao, 1997; Wilson, 2002). But, is this all of the story? From a cognitive engineering perspective, is knowledge in-the-world always to be preferred to knowledge in-the-head? From a cognitive theory perspective, does perceptual-motor activity have a privileged status over memory?

In this paper, we examine the use of knowledge in-the-world when the knowledge is on the screen, in a well-known location, and the access effort is typical of that encountered by a user of an operating system that supports multiple, overlapping windows. We argue that in routine interactive behavior, the path of least effort is a *soft constraint* that guides interactive behavior and that these soft constraints may be calculated by giving equal weight to the time required for perception, action, and memory retrieval. Milliseconds matter in that differences in effort measured in milliseconds suffice to induce users to ignore perfect knowledge in-the-world for imperfect knowledge in-the-head.

We begin by situating our account of soft constraints within the current discussion of embodied cognition. We then elaborate the concept of soft constraints by casting it in the rational analysis framework (Anderson, 1990, 1991; Oaksford & Chater, 1998). In the next section we introduce three scenarios that are typical of routine computer-human interaction and discuss the perceptual-motor and memory factors that each requires to access knowledge. We present estimates of the effort of the perceptual-motor and memory retrieval for each scenario, and use these estimates to provide a rational analysis of each scenario. We conclude the section by deriving two sets of behavioral predictions from our rational analysis; one set for the relative differences between scenarios in the number of errors made, and a second set for relative differences between scenarios in the frequency of accessing knowledge in-the-world. In Section 4, we test our predictions in two empirical studies. In Section 5, we summarize our results and discuss the implications of our work for cognitive theory and cognitive engineering.

## 2. Soft constraints in interactive behavior

For many interactive devices, the sequence and methods of operation is determined by hard constraints. For example, if your task goal is to take \$100 out of your checking account using an ATM, you must find an ATM, insert your card, key in your pin number, press fast cash, take the money, and then take the card. For any one ATM, the constraints built into its design dictate the set of possible patterns of interactive behavior (i.e., microstrategies). Knowing what hard constraints exist, it is possible to build models that bracket the minimum and maximum human performance that can be expected with a given interface design (Gray & Boehm-Davis, 2000; Kieras & Meyer, 2000).

Soft constraints complement hard constraints. Hard constraints determine what patterns of interactive behavior are possible. In contrast, rather than mandating a pattern, soft constraints suggest which of the possible patterns are likely to be chosen and executed. For example, in an empirical study reported by Gray (2000), out of nine subjects who discovered how to program a simulated VCR, seven adopted the task-to-device rule hierarchy of Fig. 1 and two adopted minor variants. In the work reported here, of the 80 subjects shown Fig. 1 as the experimenter programmed the first show, all used that task-to-device rule hierarchy to program the next four shows. Although extreme variation was possible, little variation was found. Working within the hard constraints explicitly designed into the artifact, soft constraints determined how people attempted to use the VCR.

Soft constraints arise from the rational (adaptive) nature of human cognition (Anderson, 1990; Simon, 1956); namely, that human cognition is assumed to be well adapted to the

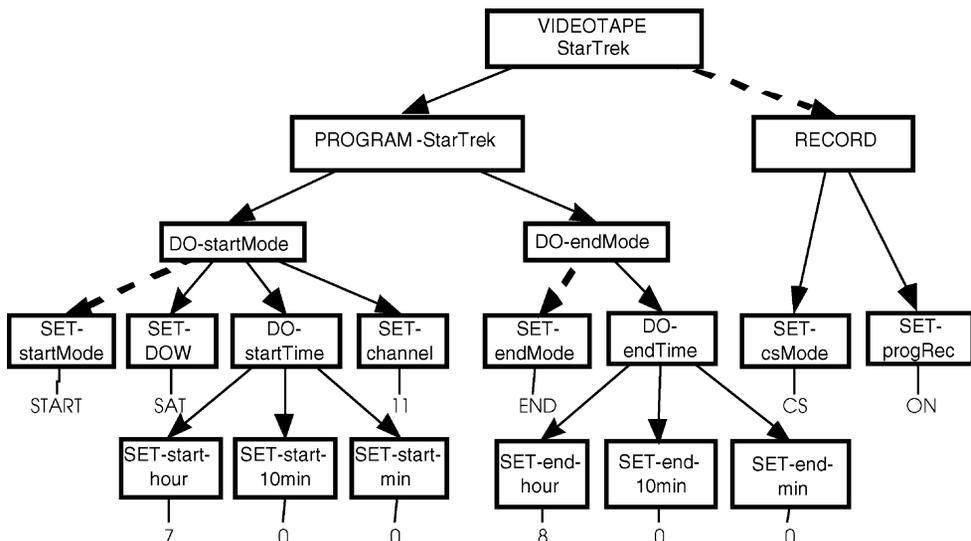


Fig. 1. The task-to-device rule hierarchy for programming the particular VCR used in experiment 1 and 2. This task-to-device rule hierarchy is largely determined by soft constraints. (Subgoals are represented by boxed nodes. Leaf nodes are unboxed and may represent multiple keystrokes. The dashed line leading from DO-startMode and DO-endMode indicate that subgoals SET-startMode and SET-endMode must be performed before the others. Contrariwise, the dashed line from VIDEOTAPE to RECORD indicates that RECORD must be performed last. With those three exceptions, the subgoals of a goal may be performed in any order.)

characteristics of the environment. Under the rationality assumption, goal-directed actions are chosen and executed through interactions between the human's adaptive mechanisms and the environment in ways that optimize efficiency and effectiveness. However, given the wide range of possible variations of natural and artificial environments, optimization can only be estimated based on local cues of the environments. Anderson's rational analysis approach (Anderson, 1990, 1991; Anderson and Schooler, 1991; Oaksford and Chater, 1998) assumes that the estimation process is based on the statistical characteristics of the natural environment. Specifically, the choice of one pattern of behavior against another depends on whether it has a higher expected utility—a measure which takes into account the statistical estimates of both effectiveness and efficiency.

We believe that the root of soft constraints lies in human adaptivity, and the measure of expected utility reflects the adaptive mechanisms in cognition. In this paper, we will use the calculations of expected utilities to make quantitative comparisons of different possible patterns of behavior. We maintain that the selection of different patterns of behavior is a non-deliberate process at the neuronal or subsymbolic level (Anderson & Lebiere, 1998).

In this paper, we ignore effectiveness to focus on how considerations of effort lead to the selection of alternative microstrategies. This focus on effort reflects a simplifying assumption that, almost by definition, common methods of interacting with standard computer interfaces accomplish what they set out to do and, hence, are highly effective. In what follows, we focus on the implications that rational analysis has for soft constraints on interactive behavior.

### 2.1. *Not all efforts are equal, are they?*

Wilson (2002) delineates six senses of the term *embodied cognition*. Of the six, the one nearest to our work is the one in which “we off-load cognitive work onto the environment” (p. 626). This sense of embodied cognition is embraced by Dana Ballard and associates (Ballard et al., 1995, 1997), who have pronounced the interactions among elementary cognitive, perceptual, and action operations at the 1/3 of a sec time scale “the embodiment level.”

Like Ballard, we see this level as where elementary operations begin to cohere into the patterns of activities that form the bases of interactive behavior. To our eye, these patterns emerge at the 1/3 of a sec level whether the analysis is done using a cognitive engineering tool such as CPM-GOMS (Gray & Boehm-Davis, 2000; Gray, John, & Atwood, 1993), ACT-R with its new buffer-based architecture (Anderson, Bothell, Byrne, & Lebiere, in press), or EPIC (Kieras & Meyer, 1997).

Where we diverge from Ballard and others is in their belief in the “minimum memory strategy” (Wilson, 2002); namely, that perceptual-motor activity has a privileged status in that, all else equal, perceptual-motor effort is generally preferred to memory effort. This belief is so pervasive that it appears to need an explicit statement to the contrary when it is not shared. For example, Shirouzu, Miyake, and Masukawa (2002) warn, in a footnote, that “We do not maintain that external resources without internal cognitive workings are preferred or that such preferences are the essential human cognitive nature” (footnote 4).

We are not the first to complain about the current unqualified enthusiasm for the utility of knowledge in-the-world. Indeed, in the context of graphical representations, Scaife and Rogers

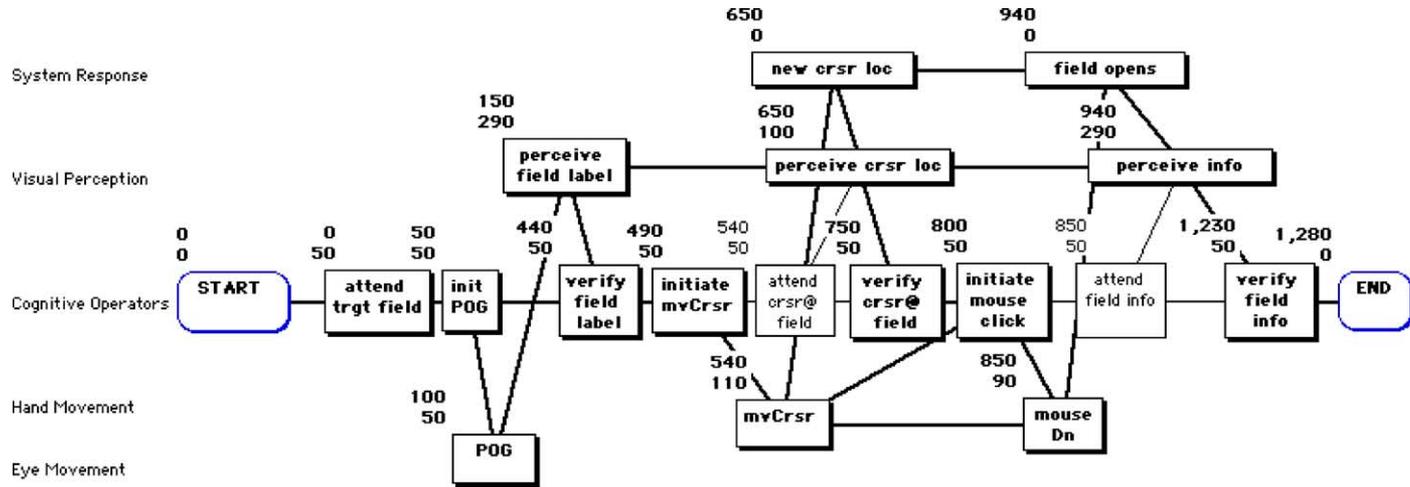


Fig. 2. A CPM-GOMS model of the microstrategy required to move to, click on, and perceive information in the Show Information Window. Total predicted time is 1,280 ms. CPM-GOMS is a network modeling technique (Schweickert et al., 2003). In the middle row are cognitive operators with a default execution time of 50-ms each. Above that line are the perceptual operators and below it are the motor operators. The flow of operators is from left-to-right with connecting lines indicating dependencies. Within an operator type the dependencies are sequential. However, between operator types the dependencies may be parallel. The numbers above each operator indicate the time, in milliseconds, for that operator to execute. Time is accumulated from left-to-right along the critical path. (The critical path is indicated by bold lines connecting shadowed boxes.) *Abbreviations:* crsr, cursor; loc, location; mvCrsr, move cursor; POG, point of gaze; trgt, target. See Gray and Boehm-Davis (2000) for more detailed information.

(1996) concluded that despite the “intuition” of the value of graphical representations “we have no well-articulated theory as to how such an advantage might work” (p. 200).

In the work presented here, we take the stance that, at the embodiment level of analysis, milliseconds matter and they matter the same regardless of the type of activity with which they are filled. Going beyond this, we sketch a process account, at the 1/3 of a sec level of analysis, that can be used to predict characteristics of task environments that facilitate or discourage the use of knowledge in-the-world for knowledge in-the-head.

## 2.2. *Patterns of interactive behavior*

Patterns of interactive behavior emerge at the embodiment level of cognition (Ballard et al., 1997). Basic patterns of activity that take about 1/3 of a sec to execute combine with other basic patterns to form microstrategies (Gray & Boehm-Davis, 2000) that accomplish a unit task (Card, Moran, & Newell, 1983).

As suggested by Fig. 2, the patterns highlight the control of interactive behavior, with central cognition orchestrating processes such as mouse movements, eye movements, perception, shifts in attention, memory encoding, and memory retrieval (memory encoding and retrieval is not shown in Fig. 2).

The time needed to execute a microstrategy is determined by the duration of operators along its critical path (Fig. 2) (Gray & Boehm-Davis, 2000; Schweickert, Fisher, & Proctor, 2003). In our analyses, no discount is given for the type of operator; that is, time for perception and motor operators that occur along the critical path is weighed the same as time for memory retrievals.

## 2.3. *Summary of soft constraints in interactive behavior*

Soft constraints bring concrete predictions based on the adaptive nature of interactive behavior. As with any rational analysis, the complete calculation of soft constraints would have to include considerations of effectiveness as well as of effort. However, to simplify our analyses we will ignore effectiveness and concentrate on effort.

In routine interactive behavior selecting a microstrategy is non-deliberate. Although selecting the best means by which to perform the next step can become the subject of deliberate selection, such deliberations would take longer than the execution of the behavior selected. Rather, the adaptive mechanisms in cognition allow non-deliberate selection of microstrategies to service our goals with a reasonable level of effectiveness and efficiency.

## 3. **Three scenarios and their perceptual-motor and memory factors for accessing knowledge**

Imagine three scenarios (each of which corresponds to an experimental condition) that vary the effort required to access knowledge in-the-world versus in-the-head:

1. Information is clearly visible on the screen in front of a user so that the user has free access to information via an eye movement (the *Free-Access* condition, see Fig. 3).

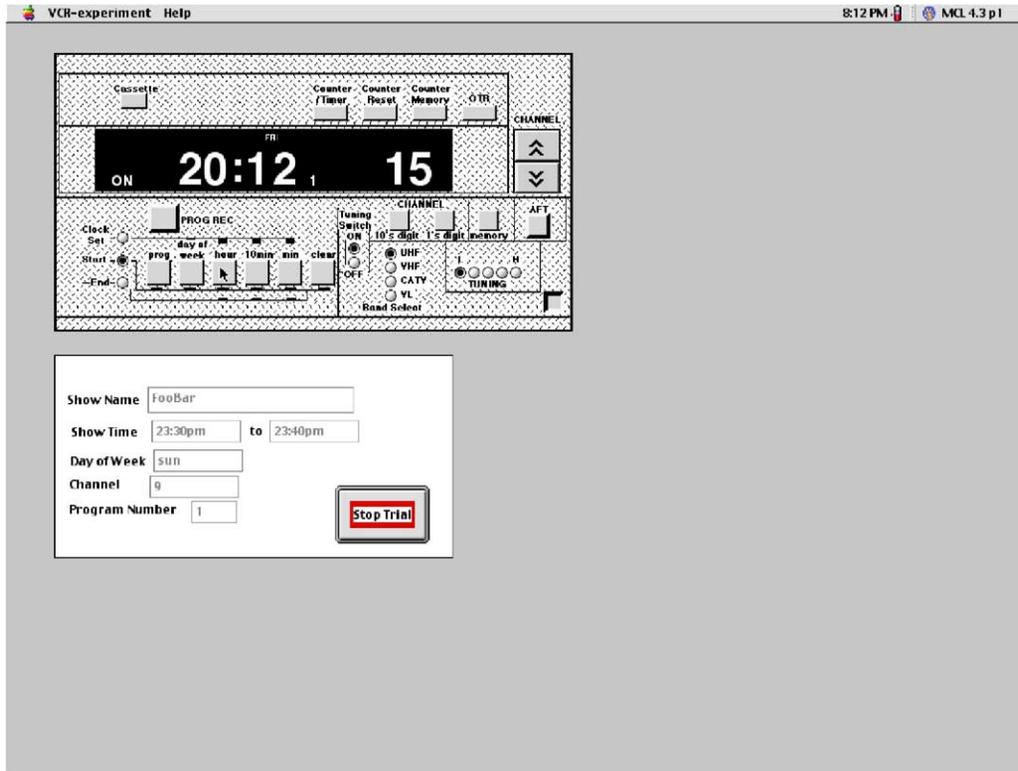


Fig. 3. Screen shot of VCR and Show Information Window for the Free-Access condition. Notice that the fields of the Show Information Window are open at all times. For the Gray-Box and Memory-Test conditions, the fields would be covered by gray boxes during the trial.

2. The window is partly visible but the desired information is covered. To uncover the information the user needs to move the mouse to and click on the window to bring it to the foreground. In experiment 1, we mimic this common circumstance by covering the fields of the information window with gray boxes. Field information is uncovered when the gray box is clicked (the *Gray-Box* condition).
3. Similar to the above, but the partly visible window contains well-learned information and bringing the window to the foreground obscures the original task window. In experiment 1, we make the material in the information window well-learned by requiring the *Memory-Test* group to study and pass a test on it before the trial begins. During programming only one window is visible at a time, either the task window or the information window.

For each scenario, perfect knowledge is readily available in-the-world. However, we predict that, between conditions, the small differences in effort required to obtain that information will lead to the occasional adoption of microstrategies that rely on error prone memory. Hence, despite perfect information in-the-world, performance will be more successful in some conditions than in others.

For these three scenarios there are four components of effort that may be required to access knowledge in-the-world or to retrieve it from in-the-head. When information is located in an uncertain location then *perceptual-motor search* is required. When information is located in a familiar location then perceptual-motor search can be replaced by *perceptual-motor access*. Alternatively, if information is to be obtained from the head, it must first be acquired thereby entailing *memory encoding* effort. Once in-the-head, using the information requires that it be retrieved from memory, thereby entailing a *memory retrieval* effort.

To simplify our analysis we will ignore two of these four effort components: perceptual-motor search and memory encoding. In our study, knowledge in-the-world is contained in an information window with well-defined fields for different categories of information (see Fig. 3). The location of the information window is constant throughout the study as is the location of the window's information fields. We assume that the effort of location learning is small (Ehret, 2000, 2002) and is incurred during the training phase of the study. Hence, during the study there is no perceptual-motor search, only perceptual-motor access.

We also ignore the effort of memory encoding. In the Memory-Test condition the information required is studied and tested before each trial begins. If a subject fails the test they must continue studying until the test is passed. Hence, in the Memory-Test condition, as in our third example, the memory encoding effort has been met before the trial begins.

For the Free-Access and Gray-Box conditions we assume that a weak memory trace is produced, one that is just strong enough to be used to program the next step in the VCR. For such a weak trace, whatever encoding occurs, can take place simultaneously with moving the eyes (Free-Access) or mouse (Gray-Box) from the information window to the VCR.

By ignoring perceptual-motor search and memory encoding, we focus our analysis on a comparison of perceptual-motor access and memory retrieval. For the Free-Access condition perceptual-motor access time includes the time needed to shift visual attention plus an eye movement to a known location and to encode the simple information at that location. For the Gray-Box and Memory-Test condition, perceptual-motor access effort also includes the time to move the cursor to and click on the gray box.

For memory retrieval, we make four assumptions, the first three of which assume the modal theory of memory contained in most introductory cognitive psychology textbooks. First, memories are retrieved in response to some environmental or mental cue. Second, cues do not always uniquely specify a single memory. Third, the order, probability of recall, and speed of recall is determined by the relative strength of the memory to the cue.

The assumption that strong memories take less time to retrieve than do weak memories is key to our arguments but is moderated by our last assumption. Fourth, memory is noisy in the signal-processing sense of the term (Altmann & Gray, 2002; Anderson & Lebiere, 1998). If a given memory,  $M$ , has a certain true strength of association,  $x$ , to a given cue then noise,  $\sigma$ , may serve to temporarily increment [ $M(x + \sigma)$ ] or decrement [ $M(x - \sigma)$ ] its true strength. This last assumption conspires to give lie to the certainty of knowledge in-the-head.

### 3.1. A rational analysis of the scenarios

The rational analysis framework leads us to make a pair of very general predictions concerning the strategies that will be used.

1. The lower the effort of perceptual-motor access relative to memory retrieval, the greater the reliance on perfect knowledge in-the-world.
2. Contrariwise, the higher the effort of perceptual-motor access relative to memory retrieval, the greater the reliance on imperfect knowledge in-the-head.

To work out the specific predictions for the three conditions of our study, Free-Access, Gray-Box, and Memory-Test, we must first come up with estimates of the efforts of perceptual-motor access versus memory retrieval for each condition. These estimates will enable us to predict the relative tradeoff of perceptual-motor access versus memory retrieval strategies between conditions. Based on these effort estimates, we will use our knowledge of cognitive science in general and memory strength in particular to derive two sets of behavioral predictions that will be tested in the experiments that follow.

### *3.1.1. Access and retrieval efforts*

For each of the three conditions, Free-Access, Gray-Box, and Memory-Test, we can derive estimates of the time needed for perceptual-motor access and memory retrieval. For the Free-Access group we use the estimate of 500-ms as the time needed to initiate an eye movement, move the eye and visual attention to a known location, perceive the simple information at that location (a single word or number), and for cognition to verify that the information had been received. This 500-ms estimate is derived from two sources. The first source is a small ACT-R 5.0 model<sup>1</sup> that simply moves visual attention from one location to another and returns the information at the second location. Using default parameters, this model yields a effort estimate (*C*) of 470 ms. The second source is the CPM-GOMS models provided by [Gray, John, & Atwood \(1993\)](#).

In terms of perceptual-motor access efforts, both the Memory-Test and Gray-Box groups have to do the same thing; namely, move visual attention, their eyes, and the cursor to the correct field in the Information Window, click on its gray box, and perceive a simple word. Extrapolating from the CPM-GOMS models and data presented by [Gray and Boehm-Davis \(2000\)](#) for the distances moved here, this time should be between 1,000 and 1,500 ms. This estimate brackets the estimate for effort obtained from two different modeling approaches. A CPM-GOMS model developed specifically for this paper (included as [Fig. 2](#)) yields an estimate of 1,280 ms; whereas a modified version of our simple ACT-R model yields an estimate for effort, *C*, of 1,120 ms. (The modifications simply added a mouse movement and click to the first ACT-R model.)

Effort, measured in time, to retrieve a weak memory is between 500 and 1,000 ms. Although for important memories, retrieval can be repeatedly cued with the same or different stimulus, 1,000 ms is the upper estimate that [Anderson and Lebiere \(1998\)](#) provide for the time that the cognitive system will allocate to any one retrieval attempt. This estimate of retrieval time for weak memories applies to the Free-Access and Gray-Box conditions.

Before the Memory-Test group begins a trial they must study and pass a test on the information required for that trial. Hence, for them, the memories retrieved are strong. Our estimate for the time needed to retrieve a strong memory ranges from 100 to 300 ms ([Altmann & Gray, 2002](#); [Byrne & Anderson, 2001](#)).

### 3.1.2. Strategic predictions by condition

The estimates of the effort derived above (and shown in Table 1) enable us to predict the relative mix of perceptual-motor access or memory retrieval for each condition. The Free-Access condition should favor perceptual-motor access over memory retrieval. Indeed, the low effort of perceptual-motor access should rule out use of the slow retrieval of weak and erroneous memories. However, although we estimated the range of memory retrieval times to be from 500 to 1,000 ms, there may well be some strong memories that can be retrieved faster than this. For the most part, most memories that can be retrieved faster than perceptual-motor access should be correct. But, because of noise in the memory system some number of erroneous traces may be relied on.

For the Gray-Box condition, the higher perceptual-motor effort should compete with the effort of retrieving a weak memory. Correct memories will tend to be weak and not much stronger than competing memories for information from older shows or other fields. Because of noise in the memory system some erroneous traces will be retrieved in place of correct traces.

Finally, for the Memory-Test condition retrieval should be much faster than perceptual-motor access making memory retrieval the preferred microstrategy. Because strong but recently rehearsed material decays rapidly (Altmann & Gray, 2002), there should be little competition from older memories with the result that the correct memories should be reliably retrieved.

The above predictions are based on characteristics of the human perceptual-motor and memory systems. Soft constraints are imposed on behavior by the interaction of these characteristics and the characteristics of the environment (i.e., in this case, the effort required to access a piece of information). The same kind of analysis can be easily carried out to derive soft constraints in other environments.

### 3.1.3. Behavioral predictions by condition

The above analyses led us to make two sets of behavioral predictions. One set is for the relative number of errors made, whereas the other set is for the relative number of times the information window is accessed.

Table 1  
Estimates (in ms) of perceptual-motor and memory retrieval effort by condition

Condition	Perceptual-motor access	Memory retrieval
Free-Access	500 <sup>a</sup>	500–1,000 <sup>b</sup> (weak)
Gray-Box	1,000–1,500 <sup>c</sup>	500–1,000 <sup>b</sup> (weak)
Memory-Test	1,000–1,500 <sup>c</sup>	100–300 <sup>d</sup> (strong)

<sup>a</sup> Estimate based on an ACT-R 5.0 model (see Note 1) and CPM-GOMS models from Gray et al. (1993).

<sup>b</sup> Estimate from Anderson and Lebiere (1998).

<sup>c</sup> Estimate based on the models developed by Gray and Boehm-Davis (2000), an ACT-R 5.0 model (see Note 1) and the CPM-GOMS model presented as Fig. 2.

<sup>d</sup> Estimates based on models developed by Altmann and Gray (2002) and Byrne and Anderson (2001).

For relative number of errors, we predict that the most errors will be made for the Gray-Box condition, followed by Free-Access, with the least being made by the Memory-Test condition. This prediction follows as the Gray-Box condition will place undue reliance on the faster retrieval of error prone weak knowledge in-the-head over the slower, but surer, perceptual-motor access of perfect knowledge in-the-world.

For its part, the Free-Access condition should rely on the combination of fast perceptual-motor access and the fast retrieval of recently encoded, but strong, memories. However, the 500 ms for perceptual-motor access means that, with a boost in activation from random noise, occasionally a weak and erroneous memory will be quickly retrieved and used in lieu of a slightly more effortful perceptual-motor access.

In contrast, as subjects in the Memory-Test condition are tested on show information before they are allowed to program a trial, this knowledge should be strong and readily accessible. We expect the least errors from this group.

We have a two-part prediction for the relative number of times the information window is accessed. The first prediction is the most obvious. If something needed for task performance is not available in-the-head then it must be acquired from in-the-world. The Memory-Test group studies and is tested on show information before each trial begins. If the memory test manipulation is successful then show information will be stored in memory and will not need to be accessed during the trial. However, neither the Free-Access nor the Gray-Box groups have the advantage of prior study. For them knowledge must be acquired from in-the-world during the trial. Hence, the Free-Access and Gray-Box groups should access show information during the trial more often than does the Memory-Test group.

On the other hand, if there is some memory then the Free-Access condition should access the Information Window more than Gray-Box because the effort of perceptual-motor access is lower. For example, assume there is a memory,  $M$ , of strength  $x$ , whose retrieval time relative to perceptual-motor (PM) access is:  $\text{TIME}[\text{PM}(\text{eye-movement})] < \text{TIME}[M(x)] < \text{TIME}[\text{PM}(\text{mouse-movement})]$ . Under these circumstances, the Free-Access condition would make an eye-movement, but rather than make a mouse-movement, the Gray-Box condition would rely on memory retrieval.

### 3.2. *Auxiliary assumption*

Our focus is on extending rational analysis to the choices made every 500–1,000 ms in the course of routine interactive behavior. The basic prediction is that non-deliberate selection acts to minimize the effort of an interaction. Our predictions depend on an auxiliary assumption that is not directly tested but will be supported if our predictions hold.

The auxiliary assumption holds that time is a reasonable surrogate for measuring cognitive as well as perceptual-motor effort. Although time may be ultimately deemed a correlate not a causal mechanism, our assumption is that for patterns of behavior lasting from 500 to 1,000 ms time is a reasonable basis for measuring effort. However, as discussed earlier, this assumption is not without its critics. Perceptual-motor effort is assumed by some to have a privileged status in that, all else equal, perceptual-motor effort is supposed to be generally preferred to memory effort. Support for our predictions will be interpreted as indirect evidence for this auxiliary assumption.

### 3.3. Summary of predictions

The predictions derived from our analysis of soft constraints in interactive behavior are clear. For errors we expect the Gray-Box condition to make the most, the Memory-Test to make the least, and the Free-Access to be somewhere in between. For the frequency of information access, we expect the Free-Access condition to access the information window more than the Gray-Box condition. Both of these conditions should access it more than the Memory-Test condition. Our claim is that the conditions we test are not extreme, but are similar to those that occur during daily use of desktop computers. Under these conditions, despite the availability of perfect knowledge in-the-world, the soft constraints analysis predicts the reliance on imperfect knowledge in-the-head.

## 4. Experimental data

In selecting a task, two criteria were important. First, we needed a clear separation between using the task interface versus accessing information for the task. Second, we wanted a task that would not force users to keep or manipulate knowledge in-the-head; that is, storage in memory for more than a few seconds should be an optional, not a necessary requirement of task performance.

These criteria led us to select the task of programming a VCR to record a television show. Meeting our first criterion, the VCR interface was constant across conditions. With the task interface held constant, we varied the ease with which information for the to-be-recorded television show (i.e., start time, end time, day-of-week, and channel) could be retrieved from memory or accessed by the perceptual-motor system. Meeting our second criterion, the VCR did not require users to keep or manipulate knowledge in-the-head. Information from the world could be obtained, used immediately, and then forgotten.

### 4.1. Experiment 1

Experiment 1 had three conditions; all used a simulation of a commercial VCR built in Macintosh Common Lisp<sup>TM</sup>.

#### 4.1.1. Methods

**4.1.1.1. Subjects.** Seventy-two undergraduates, 24 per condition, participated for course credits. Subjects were assigned to conditions randomly in blocks of threes. The experiment took approximately 30 min. Subjects were individually run.

**4.1.1.2. Data collection.** All clicks on any button object in the simulation were time stamped to the nearest tick (16.67 ms) and saved to a log file along with a complete record of the information in the VCR's displays (e.g., mode, time, day-of-week, channel, and so on). If subjects moved the cursor from the VCR window to the Show Information Window the mouse exit and mouse enter events were recorded and time stamped. Likewise, any clicks on the gray boxes covering the fields of the Show Information Window were also time stamped and saved

to the log file along with the name of the field and the information it contained. Hence, from the log file it is possible to reconstruct a complete trace of the subject's interaction with the VCR and Show Information Window.

*4.1.1.3. Procedure.* With minor differences described below, the procedure for all conditions was the same. The study began with the subject watching as the experimenter programmed the first trial of show 0. During the show 0 the task-to-device rule hierarchy shown in Fig. 1 was placed in front of the subject and the experimenter referred to it as s/he programmed the show. After the first trial the experimenter watched as the subject programmed show 0 to criterion. At that point, the experimenter removed the task-to-device rule hierarchy and left the room while the subject programmed shows 1–4. (As show 0 was an instruction and practice show, it is excluded from the analyses reported below.)

Each subject programmed shows 1–4 to the criterion of two successive correct trials. Each trial began with the subject pressing a START TRIAL button and ended with the subject pressing STOP TRIAL. At the end of each trial, the experimental software provided feedback as to how long the trial took and as to whether the show had been programmed correctly. If the show was not programmed correctly, the subject was provided feedback on the first error that the software found. The order in which errors were checked was: clock time, start time, end time, day-of-week, channel, and program record.

For all conditions and both experiments, each trial began with the VCR covered by a black box with the Show Information Window clearly visible and immediately below the VCR (see Fig. 3). In addition to fields containing the show's name, start time, end time, day-of-week, and channel, the Show Information Window also contained the START TRIAL button. Clicking on this button began the trial, changed START TRIAL to STOP TRIAL, and removed the black box that had covered the VCR.

For the Free-Access condition, the labels and fields of the Show Information Window were clearly visible throughout each trial. In contrast, for the Gray-Box condition, the labels in the Show Information Window were visible but gray boxes covered all fields except the show name. For example, to see the channel field the subject had to move the cursor to and click on the gray box covering that field. The value remained visible as long as the cursor remained in the field.

For the Memory-Test condition, clicking on the START button removed the Show Information Window and opened a memory test window. The memory test window required the subject to select the show's start-hour, start-10 min, start-min, end-hour, end-10 min, end-min, day-of-week, and channel from a series of pop-up menus. After setting the show information the subject clicked the OKAY button. If the information had not been set correctly, the subject iterated between the Show Information Window and Memory Test Window until the memory test was passed at which time they could begin the trial. (A memory test was required before each trial of each of the four shows.)

As the VCR was being programmed, we encouraged the Memory-Test group to retrieve show information from memory by discouraging the use of the Show Information Window. As per the Gray-Box condition, gray boxes covered the fields of the Show Information Window. In addition, moving the cursor out of the VCR window, caused the VCR to be covered by a black box. The black box stayed until the subject moved the cursor back to and clicked on the VCR window.

#### 4.1.2. Results and discussion

Two sets of dependent measures are analyzed. The first set contains two performance measures of errors: trials-to-criterion and a measure of goal suspension. The second set is a process measure that examines the number of times the information window was accessed.

**4.1.2.1. Performance measures: errors. Trials-to-criterion.** A trial started when the subject clicked the START TRIAL button and continued until the STOP TRIAL button was clicked. Trials for each show continued until the show was programmed correctly for two successive trials. Given that, in each condition, show information was readily available in-the-world, we might have expected all subjects in all conditions to have spent a maximum of two trials per show. A trials-to-criterion score of greater than two, reflects the number of trials that, when the subject clicked the STOP TRIAL button, were not correctly programmed. Hence, we interpret trials-to-criterion greater than two as reflecting a reliance on imperfect memory in lieu of accessing knowledge in-the-world.

A two-way analysis of variance (ANOVA) was conducted on the number of trials to reach the criterion of two successive correct shows. Condition (Free-Access, Gray-Box, Memory-Test) was a between-subjects factor and show (1–4) was within-subjects. The main effect of condition was significant,  $F(2, 69) = 4.48, p = .015$  ( $MSE = 10.04$ ), as was the main effect of show,  $F(3, 207) = 5.90, p = .0007$  ( $MSE = 5.05$ ). The interaction of condition by show was not significant ( $F < 1$ ) (see Fig. 4).

Planned comparisons by condition yielded a significant difference between Gray-Box and Memory-Test ( $p = .0002$ ) as well as between Free-Access and Memory-Test ( $p = .037$ ). The difference between the Free-Access and Gray-Box condition was not significant. Despite the ready availability of knowledge in-the-world, both the Gray-Box and Free-Access group made more errors than did the group that had show knowledge strongly encoded in-the-head.

**Goal suspension.** The trials-to-criterion measure focused our attention on the number of trials that ended in error; that is, the number of trials that ended with a show being incorrectly

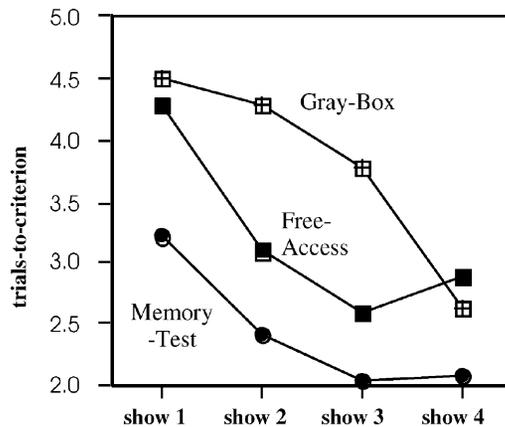


Fig. 4. Trials-to-criterion for experiment 1. Subjects were required to program each show to the criterion of two successive correct trials. Hence, for shows 3 and 4 the Memory-Test group is close to the minimum number of trials possible.

programmed. The more shows that were incorrectly programmed the greater the trials-to-criterion. In contrast, for goal suspension we examine errors that were made, but later detected and corrected on trials that ended successfully.

The measure of goal suspension is derived from Gray's (2000) goal-structure analysis of errors of performance<sup>2</sup>. For the VCR simulation there are eight fields that must be set to correctly program the VCR; day-of-week, channel, start-hr, start-10 min, start-min, end-hr, end-10 min, end-min. Given the structure of the device, the measure of goal suspension is quite simple: Once a subject starts to change a setting, how often was it abandoned before being correctly completed? For example, if for show 2 the to-be-set channel was 21, and the current channel was 11, then if the subject began setting the channel but stopped before 21 (e.g., going off to set the day-of-week), then this is one goal suspension.

For goal suspensions, we examined only trials that were successfully programmed. In the context of a successfully programmed trial, goal suspensions are potential errors. They require that the subject detect that the setting is not complete and correct the setting before pressing the STOP TRIAL button. Note that accessing show information during a setting was not considered goal suspension. For example, if a subject started programming the channel setting, interrupted his or herself to check the Show Information Window, and then resumed programming the channel—this would not be considered a goal suspension. We interpret goal suspensions as due to reliance, at least temporarily, on imperfect knowledge in-the-head rather than on perfect knowledge in-the-world. If subjects compare the current setting of, e.g., channel, with the value of channel in the Show Information Window then they would not stop, but would continue programming until the current channel matched the goal channel.

Goal suspensions are a rarity. Examining patterns of goal suspensions requires that a vast quantity of correct data be collected and parsed. Across all three conditions of experiment 1, 36,877 mouse clicks were collected and time stamped on correct trials. These mouse clicks were parsed into 12,560 goals using the action-protocol analyzer developed by Fu (2001) with the task-to-device rule hierarchy shown in Fig. 1. Less than 1% of all goals, 122, resulted in goal suspensions. For each group the mean number of goal suspensions per subject is shown in Fig. 5.

An overall ANOVA produced a marginally significant effect,  $F(2, 69) = 2.64$ ,  $p = .078$ . The statistical significance bars (SSBs) in Fig. 5 are based on planned comparisons. If two SSBs

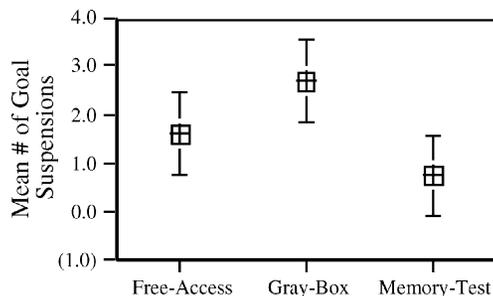


Fig. 5. Mean goal suspensions per subject across the three conditions. statistical significance bars (SSBs) show the pairwise statistical significance between means.

look different (i.e., they do not overlap), the corresponding pairwise comparison is different (at the .05 level of significance adopted for this study) (for more information on SSBs see Schunn, 2000). As indicated by Fig. 5, the Gray-Box condition made significantly more goal suspensions than did the Memory-Test condition, but there were no significant differences between the other comparisons.

A  $\chi^2$  comparison that looked at whether or not each subject made goal suspensions, was significant ( $p = .05$ ). Fifty percent of the Free-Access subjects made goal suspensions, 75% of the Gray-Box subjects, and 42% of the Memory-Test subjects.

*4.1.2.2. Discussion of performance measures.* The first two dependent measures, trials-to-criterion and goal suspensions, yield a consistent pattern. The Memory-Test condition is best, and the Gray-Box is worst with the Free-Access condition somewhere in the middle. These data present us with an interesting quandary. All groups had access to all show information at all times, yet they made errors that kept them in the study longer than they needed to be.

The subjects had to program each show until they got it correct twice in succession. Hence, the penalty for ending the trial in error was having to stay in the experiment longer. Subjects in the Free-Access or Gray-Box groups could have matched the performance of the Memory-Test group by simply comparing their settings against the Show Information Window before clicking the STOP TRIAL button. Similarly, the penalty for a goal suspension was having to go back and complete the suspended goal at a later time at the risk of ending the trial in error. Subjects in the Gray-Box condition could have easily double-checked show information before suspending their current goal. Both of these measures, trials-to-criterion and goal suspensions, suggest that soft constraints lead to reliance on imperfect memory for show information rather than more reliable perceptual-motor access.

*4.1.2.3. Process measure: accesses of knowledge in-the-world.* Our process measure counts the number and the pattern of information accesses to the Show Information Window. For the Memory-Test and Gray-Box conditions, each click on a gray box was counted. The pattern of when information was accessed versus when the information was programmed was derived from the log files.

This process measure can be used to address two questions. The first is a construct validity issue (Gray & Salzman, 1998): Did the Memory-Test manipulation lead to the retrieval of show information from memory instead of accessing it from the display? The second examines what the patterns of information access reveal about the use of knowledge in-the-world.

*Construct validity.* Did the Memory-Test group rely on memory retrieval or on perceptual-motor access? Throughout shows 1–4, the 24 subjects in the Gray-Box condition clicked on information fields 293 times over 223 correct trials for an average of 1.31 checks per show. In contrast, the 24 subjects in the Memory-Test condition clicked on an information field 10 times during 205 correct trials for an average of 0.05 checks per show. This contrast suggests that the memory manipulation was successful and that the Memory-Test group almost exclusively relied on retrievals from memory as their source of show information.

*Patterns of information access.* Given that subjects in the Gray-Box condition could access knowledge in-the-world whenever they wanted it, can their patterns of information access provide any clue regarding why this group did not do as well as the Memory-Test group?

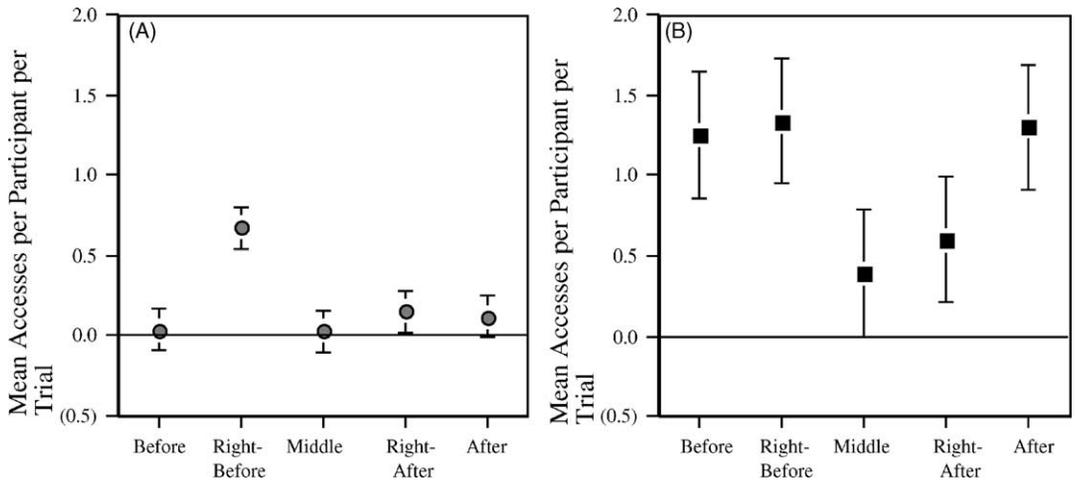


Fig. 6. (A) For the Gray-Box condition from experiment 1, the graph shows the mean accesses per subject per trial, SSBs, based on 24 subjects, show the pairwise statistical significance between means. (B) For the Free-Access condition from experiment 2, the graph shows the mean accesses per subject per trial, SSBs, based on eight subjects, show the pairwise statistical significance between means.

Fig. 6A shows the mean number of information accesses per correct trial per subject for the Gray-Box condition. Each information access was categorized by when it occurred in relation to when the information was used. For example, if a subject accessed channel information but set something else before setting channel, this access was classified as *before*. If after accessing the channel information the subject's next act was to program the channel setting, this access was classified as *right-before*. Any interruption of a setting to access the information for that setting was classified as *middle*. If immediately after setting the channel the subject's next act was to access the channel information, this access was classified as *right-after*. Any later access of an information field was classified as *after*.

Of these five categories of information access we will be most interested in the *right-before*, *middle*, and *right-after* categories. The *right-before* and *right-after* categories refer to well-defined points in time; immediately before or after a setting was programmed. The *middle* category refers to a well-defined time interval. In contrast, the *before* category refers to the time from the beginning of the trial until immediately after the prior setting was programmed. Complementary, the *after* category refers to a time after programming had begun on the next setting to the end of the trial. Hence, in contrast to *right-before*, *middle*, and *right-after*, the *before* and *after* categories refer to events that could occur at any place within an uncertain time interval.

A within-subject ANOVA yielded significant between-category differences in when the Gray-Box group accessed show information,  $F(4, 92) = 15.36$ ,  $p < .0001$  ( $MSE = 0.11$ ). The SSBs in Fig. 6A are based on the Tukey Honestly Significant Difference (HSD) test. As shown by the SSBs, more accesses were performed *right-before* the information was needed than at any other time. There were no significant pairwise comparisons between any of the other access categories.

Occasionally, the Gray-Box group will access knowledge in-the-world right before they program the setting. However, they are unlikely to access this knowledge while they are programming a setting (middle), and they are equally unlikely to access it right after they have programmed a setting. Any verifications of the correctness of their settings that the Gray-Box group is doing must be based on retrieval of imperfect knowledge in-the-head, not on a comparison with perfect knowledge in-the-world.

We will return to a discussion of the experiment 1 information access data after presenting experiment 2.

## 4.2. Experiment 2

Experiment 1 was interesting but incomplete, as it provided no information on how often or when the Free-Access condition accessed show information. To remedy this deficit we conducted experiment 2.

Experiment 2 was run to provide eye-tracking data on the Free-Access condition. With these data, we could examine the frequency and patterns with which the Free-Access group accessed the fields in the Show Information Window.

### 4.2.1. Method

Experiment 2 had one condition that replicated the Free-Access condition with one main difference: subjects were eye-tracked as they programmed the VCR. To facilitate eye-tracking, the size of the Show Information Window was increased to increase the visual separation between each of the information fields.

*4.2.1.1. Subjects.* We report results from the first eight undergraduates who we could successfully eye track. All subjects, whether or not they could be eye-tracked, received course credit for their participation. Because of the necessity to calibrate the eye-tracker on each subject, experiment 2 took approximately 45 min.

*4.2.1.2. Eye tracking.* Eye tracking was performed using an ASL 504 remote optics eye tracker. Head movements were tracked using a Flock-of-Birds<sup>TM</sup> magnetic head tracker. Eye data was sampled and saved to a log file 60 times per second (once every 16.67 ms).

Fixations were determined using the algorithm developed by Karsh and Breitenbach (1983). Basically, we say that a fixation occurs when at least six consecutive data points fall within a  $3 \times 4$  pixel rectangle (where the definition of “consecutive” points is that they have to be less than 32 ms apart). Areas of interest were created around each information box. Consecutive fixations in the same area of interest were counted as a single access.

### 4.2.2. Results

With fewer subjects, 8 versus 24, the variability for the experiment 2 Free-Access group was greater than that for the comparable experiment 1 group. However, performance on trials-to-criterion and goal suspensions were within the range we would expect based on the experiment 1 data.

Eye tracking yields a much finer grain of analysis and greater data density than the measures we reported in experiment 1. Hence, despite the fewer number of subjects, the greater data density per subject might be expected to yield relatively stable estimates of the pattern of information access.

Fig. 6B shows the mean number of information accesses per correct trial per subject for the Free-Access condition in experiment 2. A within-subject ANOVA showed the between category differences to be significant,  $F(4, 28) = 5.38$ ,  $p = .002$ ,  $MSE = 0.29$ . The SSBs in Fig. 6B are based on the Tukey HSD test. The SSBs show no difference in number of accesses between the before, right-before, and after categories. However, each of these three categories significantly differs from the middle category and is marginally different from the right-after category. There are no differences between the middle and right-after categories.

#### 4.3. Discussion of information access in experiments 1 and 2

The patterns of information access in the Gray-Box condition of experiment 1 and the Free-Access condition of experiment 2 have some interesting similarities. Both groups are more likely to access information right-before they need it instead of when they are using it (i.e., middle) or right-after. Apparently, both groups were so complacent in their ability to retrieve the correct information from memory or in the feeling-of-knowing that came from looking at a value they had just set that they were unwilling to pay the perceptual-motor effort needed to verify that the current setting was, indeed, the target setting.

The differences in patterns of information access are as revealing as the similarities. First, the lower the perceptual-motor effort required to access information, the more frequent the accesses. Over all categories the experiment 2 Free-Access group is 4.3 times more likely to access information than is the Gray-Box group. However, although the number of accesses decreases in all categories between the Free-Access and Gray-Box conditions, the one category that is partially protected is the right-before category. The Gray-Box group appears to devote a higher proportion of its information accesses to the right-before category than does the Free-Access group. The higher number of accesses before and after suggests that the Free-Access group does more advance storage than the Gray-Box group and more comparing of the VCR settings to the show information.

## 5. General discussion

### 5.1. Success of predictions—support for hypotheses and assumptions

Plugging Table 1 estimates of perceptual-motor and cognitive effort into our rational analysis led us to derive soft constraints in different experimental conditions and make several predictions concerning performance and process. For performance we predicted that the Memory-Test condition would be best and Gray-Box the worst with the Free-Access group somewhere in the middle. We found this rank-order with both performance measures. For trials-to-criterion there were significant differences between Memory-Test and each of the other two conditions, but not significant differences between these two. For goal suspensions planned comparisons

revealed a significant difference between Memory-Test and Gray-Box with no other significant between group differences.

The process measure supported the assumption that the Memory-Test group would rely on memory retrieval rather than perceptual-motor access. Eyeballing the difference in this measure across experiments suggested that the Free-Access group made many more perceptual-motor accesses of the Information Window than did the Gray-Box group. This difference supports the prediction that information accesses would vary as a function of perceptual-motor effort.

The process measure directly supports the prediction that on many occasions when the Gray-Box or Free-Access condition might have been expected to rely on knowledge in-the-world they relied, instead, on knowledge in-the-head. For example, while programming (middle) a setting and right-after programming a setting the Gray-Box condition made almost no accesses of the Show Information Window (see Fig. 6A). The Free-Access condition made only slightly more middle and right-after accesses, averaging approximately 1/2 an access per person per trial (see Fig. 6B).

These behavioral measures support our predictions based on our notion of least effort tradeoffs as soft constraints that govern non-deliberate selection of microstrategies in interactive behavior. The success of our analysis in predicting experimental outcomes provides indirect support for our auxiliary assumption. For patterns of interactive behavior lasting between 500 and 1,000 ms, time is a reasonable surrogate measure for cognitive as well as for perceptual-motor effort. Unless our estimates of the duration of memory and perceptual-motor processes are way off, the success of our predictions suggests that the time spent retrieving something from memory is weighed the same as time spent in perceptual-motor activity.

## 5.2. Implications

The perceptual-motor and memory effort manipulated in these studies are of the same order of magnitude as the effort paid by the typical user of direct-manipulation interfaces. The effort associated with the Memory-Test condition is similar to that paid by the author who relies on his strong memory for details contained in a chart tucked away at the end of the manuscript. The Free-Access condition is similar in perceptual-motor effort to many situations in which information that is available in one open window is required by a program running in another open window. Finally, subjects in the Gray-Box condition spent an effort equivalent to that required by users who must move to and click on a partially covered window to bring the information it contains to the foreground. Indeed, given the pedestrian nature of the manipulations, it is interesting and important that the three conditions produced the pattern of results that they did.

It may not be completely surprising that such small differences in perceptual-motor efforts affected strategy (Wickens, 1992), but we are surprised that they influenced performance. Indeed, the most striking aspect of the between-group differences in performance is that all were avoidable. All performance differences can be traced to differences in willingness to either memorize or access show information. For each trial the Memory-Test group had quick and reliable access to show information in memory. The other groups made more errors that resulted in more trials-to-criterion and more goal-suspensions. Apparently, verification is lower

effort—and hence more likely—if based on knowledge in-the-head rather than accessing knowledge in-the-world.

The results we obtained demonstrate the “bounds” of rationality (Simon, 1956). Since the adaptive mechanisms in cognition operate on local estimates of the characteristics of the environment, there is no guarantee that the pattern of behavior chosen will lead to global optimal performance. However, these bounds bring out the importance of soft constraints; expectations of global optimal performance are naïve unless soft constraints are engineered to support it.

### *5.2.1. Implications for routine interactive behavior*

Although the level of perceptual-motor and memory effort manipulated in these studies was representative of that encountered in many human-system interactions, this effort is much lower than that involved in many others. For example, for a typical process control operator, accessing knowledge in-the-world may require more than simply clicking on a gray box, it may require getting out of a seat and moving across the room to an information display. Similarly, unlike the situations we studied, visually busy web pages impose a substantial search effort on accessing knowledge in-the-world. Hence, for much routine interactive behavior, we would expect the effort-benefit tradeoff to favor accessing imperfect knowledge in-the-head over perfect knowledge in-the-world.

### *5.2.2. Implications for design*

Some readers may object, as did one reviewer of our conference presentation (Gray & Fu, 2001), that if the VCR had been designed differently then the observed failures to access knowledge in-the-world would not have occurred. However, this observation is not an objection to the current research but, rather, is precisely the point. It is well-established that design of the task environment influences the strategies adopted (Cary & Carlson, 1999, 2001; Neth & Payne, 2001; O’Hara & Payne, 1998, 1999; Payne, Howes, & Reader, 2001). The goal of our research is to understand how interactive behavior emerges from the constraints and opportunities provided by the interaction of embodied cognition with the task being performed and the interface designed to perform the task. The difficulty lies in understanding how small changes in interface design interact with embodied cognition to produce interactive behavior. Hence, the proper focus of our study is not the interface per se, but the human. What is important is not the observation that different interface designs produce different patterns of interactive behavior, but understanding the interaction of design with embodied cognition that leads to these different patterns.

### *5.2.3. Implications for embodied cognition*

Our results would be different if the knowledge in-the-world we studied was not text and linguistic but was something else. Our results would be different if the amount of information required for task performance was greater or lesser than what we used. Our title is misleading as once information is acquired via the perceptual-motor system it is then in-the-head not in-the-world.

We have heard and agree to all of these points. Indeed, our point is precisely that the most adaptive pattern of interactive behavior is one that is the least effort given the current task and the current task environment. Change the environment or the task then another pattern will be most adaptive. However, like Scaife and Rogers (1996), rather than appealing to hand-waving

theories or intuitive accounts of what should be easier for people to do, we call for an account of the control of interactive behavior; that is, the ways in which central cognition orchestrates processes such as mouse movements, eye movements, perception, shifts in attention, memory encoding, and memory retrieval.

The contrary appeal to intuitive accounts seems to come from both the practitioner and researcher communities. The practitioners seize on their bumper sticker approaches to design guidance (put knowledge in-the-world) while the researchers have theirs (“we off-load cognitive work onto the environment,” Wilson, 2002).

We argue for a more nuanced approach, one that does not presume the privileged status of any location or type of operation. Indeed, we would rephrase Wilson’s third sense of embodied cognition to read, “the cognitive control of interactive behavior minimizes effort by using a least effort combination of the mechanisms available to it.” All mechanisms or subsystems are on the table. There is no reason to think that one mechanism or subsystem has a privileged status in relation to another.

The leap from “we can use knowledge in-the-world” to “knowledge in-the-world has a privileged status compared to knowledge in-the-head” is not required. What is required is a careful and fine-grained analysis of the patterns of interactive behavior needed to perform a task.

## 6. Conclusions: soft constraints in interactive behavior

It is not surprising that people who are forced to memorize show information do well. Nor would it be surprising to find that people who must acquire information from their environment take longer and require more steps than those who have already acquired it. What is surprising is that perfect knowledge in-the-world produces less than perfect performance even in a simple task whose demands are about what we all encounter daily in our use of interactive systems.

Soft constraints are imposed by the designer on the user. Whether or not designers are aware of soft constraints or of their effect on user performance, soft constraints exist and their influence is real. There is a clear need for the cognitive engineering community to develop tools and a new generation of analytic guidelines that can build a consideration of soft constraints into artifact design and facilitate the evaluation of soft constraints after an artifact has been built.

Our analysis of soft constraints is based on process models at the embodiment level of analysis. It is at this level that we see cognitive, perceptual, and action operators orchestrated into patterns of interactive behavior. These patterns form the activities and microstrategies of embodied cognition.

This convergence of theory and practice suggests that the embodiment level is the right level of description for functional cognition. We are optimistic that a cognitive science that is inspired by problems in the world and attempts to develop engineering tools for these problems will be grounded at the embodiment level.

## Notes

1. A compressed file (StuffIt™ format) containing two Lisp files and an Excel workbook may be downloaded from the annex maintained by the *Cognitive Science Society*,

see supplementary material <http://cogsci.psy.utexas.edu/supplements/>. The Lisp files are written in Macintosh Common Lisp (MCL) but use the uniform-interface standard for ACT-R and hence the windows created as well as the models should run under ACL Lisp (though this has not been tested). Each Lisp file implements one of the models discussed in this section. The Excel workbook includes sample runs of each model and estimates of effort after 1, 10, 100, and 1,000 cycles.

- Note that the terminology has changed from that used in Gray (2000). What we are referring to as *goal suspensions* were referred to there as *premature pops*.

## Acknowledgments

The work reported was supported by a grants from the Air Force Office of Scientific Research AFOSR# F49620-97-1-0353 and F49620-03-1-0143, as well as the Office of Naval Research ONR# N000140310046.

## References

- Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*(1), 27–33.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences, 14*(3), 471–517.
- Anderson, J. R., Bothell, D., Byrne, M. D., & Lebiere, C. (in press). An integrated theory of the mind. *Psychological Review*.
- Anderson, J. R., & Lebiere, C. (Eds.). (1998). *Atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*, 396–408.
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*(1), 66–80.
- Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*(4), 723–742.
- Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review, 108*(4), 847–869.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human–computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cary, M., & Carlson, R. A. (1999). External support and the development of problem-solving routines. *Journal of Experimental Psychology-Learning Memory and Cognition, 25*(4), 1053–1070.
- Cary, M., & Carlson, R. A. (2001). Distributing working memory resources during problem solving. *Journal of Experimental Psychology-Learning Memory and Cognition, 27*(3), 836–848.
- Ehret, B. D. (2000). Learning where to look: The acquisition of location knowledge in display-based interaction. *Dissertation Abstracts International: Section B: The Sciences & Engineering, 60*(10-B), 5239.
- Ehret, B. D. (2002). Learning where to look: Location learning in graphical user interfaces. *CHI Letters, 4*(1), 211–218.
- Frohlich, D. M. (1997). Direct manipulation and other lessons. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of human–computer interaction* (2nd ed., pp. 463–488). New York: Elsevier.
- Fu, W.-T. (2001). ACT-PRO: Action protocol tracer—a tool for analyzing discrete action protocols. *Behavior Research Methods, Instruments, & Computers, 33*(2), 149–158.

- Gray, W. D. (2000). The nature and processing of errors in interactive behavior. *Cognitive Science*, 24(2), 205–248.
- Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied*, 6(4), 322–335.
- Gray, W. D., & Fu, W.-T. (2001). Ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head: Implications of rational analysis for interface design. *CHI Letters*, 3(1), 112–119.
- Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world performance. *Human-Computer Interaction*, 8(3), 237–309.
- Gray, W. D., & Salzman, M. C. (1998). Damaged merchandise? A review of experiments that compare usability evaluation methods. *Human-Computer Interaction*, 13(3), 203–261.
- Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human-Computer Interaction*, 1(4), 311–338.
- Karsh, R., & Breitenbach, F. W. (1983). Looking at looking: The amorphous fixation measure. In R. Groner, C. Menz, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and psychological functions: International views* (pp. 53–64). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12(4), 391–438.
- Kieras, D. E., & Meyer, D. E. (2000). The role of cognitive task analysis in the application of predictive models of human performance. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive task analysis* (pp. 237–260). Mahwah, NJ: Lawrence Erlbaum Associates.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–99.
- Neth, H., & Payne, S. J. (2001). Addition as interactive problem solving. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 698–703). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Norman, D. A. (1989). *The design of everyday things*. New York: DoubleDay.
- O'Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive Psychology*, 35, 34–70.
- O'Hara, K. P., & Payne, S. J. (1999). Planning and the user interface: The effects of lockout time and error recovery cost. *International Journal of Human-Computer Studies*, 50(1), 41–59.
- Oaksford, M., & Chater, N. (Eds.). (1998). *Rational models of cognition*. New York: Oxford University Press.
- Payne, S. J., Howes, A., & Reader, W. R. (2001). Adaptively distributing cognition: A decision-making perspective on human-computer interaction. *Behaviour & Information Technology*, 20(5), 339–346.
- Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, 45(2), 185–213.
- Schunn, C. D. (2000). Statistical significance bars (SSB): A way to make graphs more interpretable. Retrieved February 08, 2004, from <http://www.Irdc.pitt.edu/schunn/SSB/index.html>.
- Schweickert, R., Fisher, D. L., & Proctor, R. W. (2003). Steps toward building mathematical and computer models from cognitive task analyses. *Human Factors*, 45(1), 77–103.
- Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science*, 26(4), 469–501.
- Shneiderman, B. (1982). The future of interactive systems and the emergence of direct manipulation. *Behaviour and Information Technology*, 1(3), 237–256.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.
- Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636.