



ELSEVIER

Cognitive Science xxx (2004) xxx–xxx

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

5

6           Soft constraints in interactive behavior: the case of

7           ignoring perfect knowledge in-the-world for

8           imperfect knowledge in-the-head<sup>☆</sup>

9           Wayne D. Gray<sup>a,\*</sup>, Wai-Tat Fu<sup>b</sup>

10           <sup>a</sup>*Rensselaer Polytechnic Institute, Carnegie Building (Room 108), 110 8th Street, Troy, NY 12180, USA*

11           <sup>b</sup>*Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA*

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

13   **Abstract**

14    Constraints and dependencies among the elements of embodied cognition form patterns or micro-

15    strategies of interactive behavior. Hard constraints determine which microstrategies are possible. Soft

16    constraints determine which of the possible microstrategies are most likely to be selected. When selection

17    is non-deliberate or automatic the least effort microstrategy is chosen. In calculating the effort required to

18    execute a microstrategy each of the three types of operations, memory retrieval, perception, and action,

19    are given equal weight; that is, perceptual-motor activity does not have a privileged status with respect to

20    memory. Soft constraints can work contrary to the designer's intentions by making the access of perfect

21    knowledge in-the-world more effortful than the access of imperfect knowledge in-the-head. These impli-

22    cations of soft constraints are tested in two experiments. In experiment 1 we varied the perceptual-motor

23    effort of accessing knowledge in-the-world as well as the effort of retrieving items from memory. In

24    experiment 2 we replicated one of the experiment 1 conditions to collect eye movement data. The results

25    suggest that milliseconds matter. Soft constraints lead to a reliance on knowledge in-the-head even when

26    the absolute difference in perceptual-motor versus memory retrieval effort is small, and even when relying

27    on memory leads to a higher error rate and lower performance. We discuss the implications of soft con-

28    straints for routine interactive behavior, accounts of embodied cognition, and tool and interface design.

29    © 2004 Cognitive Science Society, Inc. All rights reserved.

30    *Keywords:* Interactive behavior; Rational analysis; Knowledge in-the-world; Human–computer

31    interaction; Embodied cognition; Embodiment level; Soft constraints; Bounded rationality; Interface

32    design; Information access; Eye tracking

☆ 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. Tel.: +1-518-276-3315; fax: +1-518-276-8268.

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

## 33 1. Introduction

34 Knowledge can be in-the-world or in-the-head (see, e.g., Larkin & Simon, 1987; Norman,  
35 1989). A well-designed interface can place knowledge in-the-world so that it is available in  
36 a known location when a user needs it. Support for acquiring knowledge from in-the-world  
37 rather than retrieving it from in-the-head is widely touted as one of the main advantages of di-  
38 rect manipulation interfaces over command language ones (Frohlich, 1997; Hutchins, Hollan,  
39 & Norman, 1985; Shneiderman, 1982). This cognitive engineering view is congruent with  
40 accounts of embodied cognition that suggest a privileged status for perceptual-motor ef-  
41 fort compared to retrieval effort (Ballard, Hayhoe, & Pelz, 1995; Ballard, Hayhoe, Pook,  
42 & Rao, 1997; Wilson, 2002). But, is this all of the story? From a cognitive engineering per-  
43 spective, is knowledge in-the-world always to be preferred to knowledge in-the-head? From  
44 a cognitive theory perspective, does perceptual-motor activity have a privileged status over  
45 memory?

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

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

## 66 2. Soft constraints in interactive behavior

67 For many interactive devices, the sequence and methods of operation is determined by hard  
68 constraints. For example, if your task goal is to take \$100 out of your checking account using  
69 an ATM, you must find an ATM, insert your card, key in your pin number, press fast cash, take  
70 the money, and then take the card. For any one ATM, the constraints built into its design dictate  
71 the set of possible patterns of interactive behavior (i.e., microstrategies). Knowing what hard  
72 constraints exist, it is possible to build models that bracket the minimum and maximum human

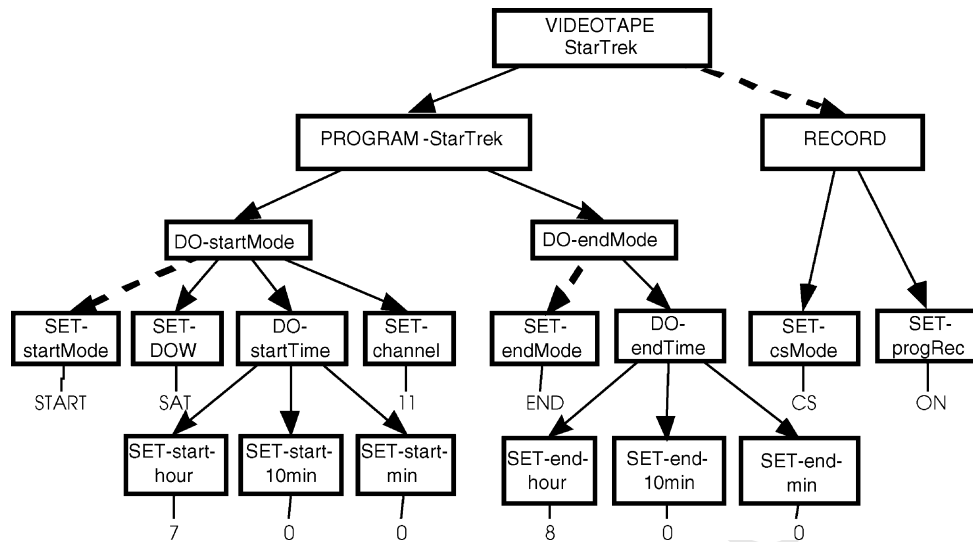


Fig. 1. The task-to-device rule hierarchy for programming the particular VCR used in experiment 1 and 2. This task-to-device rule hierarchy is largely determined by soft constraints. (Subgoals are represented by boxed nodes. Leaf nodes are unboxed and may represent multiple keystrokes. The dashed line 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.)

73 performance that can be expected with a given interface design (Gray & Boehm-Davis, 2000;  
74 Kieras & Meyer, 2000).

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

85 Soft constraints arise from the rational (adaptive) nature of human cognition (Anderson,  
86 1990; Simon, 1956); namely, that human cognition is assumed to be well adapted to the  
87 characteristics of the environment. Under the rationality assumption, goal-directed actions are  
88 chosen and executed through interactions between the human's adaptive mechanisms and the  
89 environment in ways that optimize efficiency and effectiveness. However, given the wide range  
90 of possible variations of natural and artificial environments, optimization can only be estimated  
91 based on local cues of the environments. The rational analysis by Anderson (1990, 1991),  
92 Anderson and Schooler (1991), and Oaksford and Chater (1998) assumes that the estimation  
93 process is based on the statistical characteristics of the natural environment. Specifically, the

94 choice of one pattern of behavior against another depends on whether it has a higher expected  
95 utility—a measure which takes into account the statistical estimates of both effectiveness and  
96 efficiency.

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

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

#### 108 2.1. *Not all efforts are equal, are they?*

109 Wilson (2002) delineates six senses of the term *embodied cognition*. Of the six, the one near-  
110 est to our work is the one in which “we off-load cognitive work onto the environment” (p. 626).  
111 This sense of embodied cognition is embraced by Dana Ballard and associates (Ballard et al.,  
112 1995, 1997), who have pronounced the interactions among elementary cognitive, perceptual,  
113 and action operations at the 1/3 of a sec time scale “the embodiment level.”

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

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

127 We are not the first to complain about the current unqualified enthusiasm for the utility of  
128 knowledge in-the-world. Indeed, in the context of graphical representations, Scaife and Rogers  
129 (1996) concluded that despite the “intuition” of the value of graphical representations “we have  
130 no well-articulated theory as to how such an advantage might work” (p. 200).

131 In the work presented here, we take the stance that, at the embodiment level of analysis,  
132 milliseconds matter and they matter the same regardless of the type of activity with which they  
133 are filled. Going beyond this, we sketch a process account, at the 1/3 of a sec level of analysis,  
134 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.

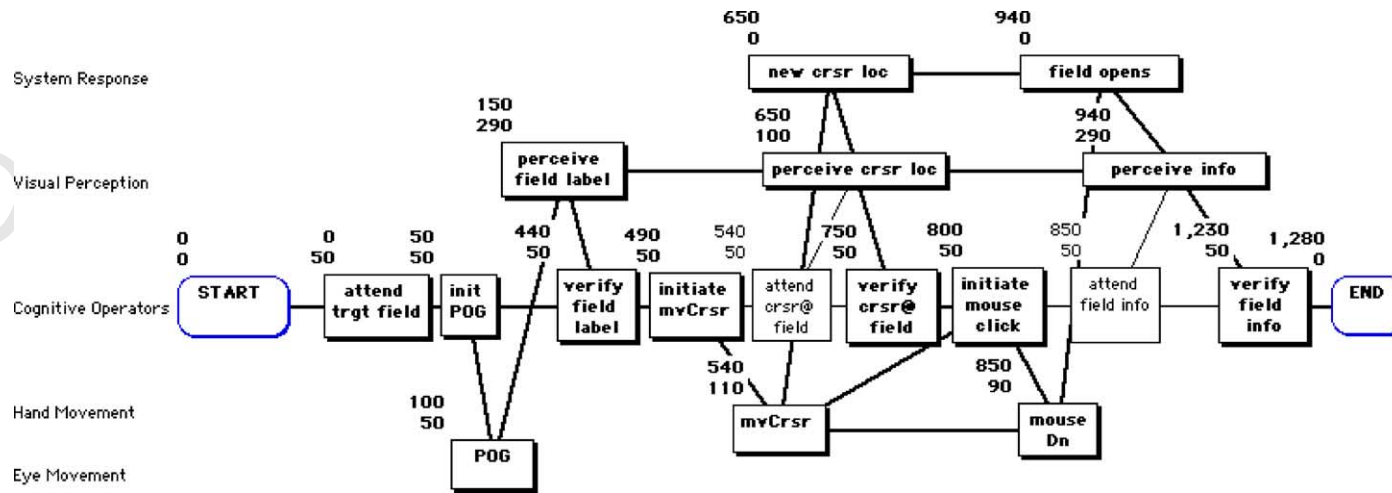


Fig. 2. A CPM-GOMS model of the microstrategy required to move to, click on, and perceive information in the Show Information Window. Total predicted time is 1,280 ms. CPM-GOMS is a network modeling technique (Schweickert et al., 2003). In the middle row are cognitive operators with a default execution 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.

## 135 2.2. *Patterns of interactive behavior*

136 Patterns of interactive behavior emerge at the embodiment level of cognition (Ballard et al.,  
137 1997). Basic patterns of activity that take about 1/3 of a sec to execute combine with other  
138 basic patterns to form microstrategies (Gray & Boehm-Davis, 2000) that accomplish a unit  
139 task (Card, Moran, & Newell, 1983).

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

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

## 149 2.3. *Summary of soft constraints in interactive behavior*

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

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

## 159 **3. Three scenarios and their perceptual-motor and memory factors for accessing** 160 **knowledge**

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

- 163 1. Information is clearly visible on the screen in front of a user so that the user has free  
164 access to information via an eye movement (the *Free-Access* condition, see Fig. 3).
- 165 2. The window is partly visible but the desired information is covered. To uncover the  
166 information the user needs to move the mouse to and click on the window to bring it to  
167 the foreground. In experiment 1, we mimic this common circumstance by covering the  
168 fields of the information window with gray boxes. Field information is uncovered when  
169 the gray box is clicked (the *Gray-Box* condition).
- 170 3. Similar to the above, but the partly visible window contains well-learned information  
171 and bringing the window to the foreground obscures the original task window. In exper-

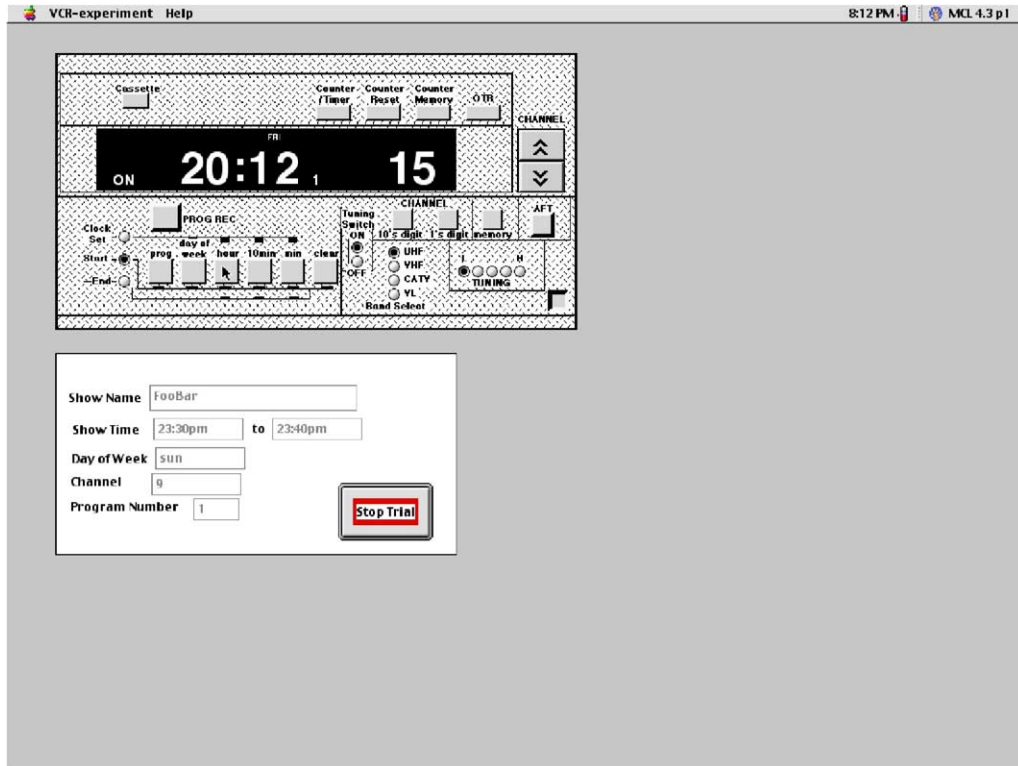


Fig. 3. Screen shot of VCR and Show Information Window for the Free-Access condition. Notice that the fields of the Show Information Window are open at all times. For the Gray-Box and Memory-Test conditions, the fields would be covered by gray boxes during the trial.

172            ment 1, we make the material in the information window well-learned by requiring the  
 173            *Memory-Test* group to study and pass a test on it before the trial begins. During program-  
 174            ming only one window is visible at a time, either the task window or the information  
 175            window.

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

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

188 To simplify our analysis we will ignore two of these four effort components: perceptual-motor  
189 search and memory encoding. In our study, knowledge in-the-world is contained in an infor-  
190 mation window with well-defined fields for different categories of information (see Fig. 3).  
191 The location of the information window is constant throughout the study as is the location of  
192 the window's information fields. We assume that the effort of location learning is small (Ehret,  
193 2000, 2002) and is incurred during the training phase of the study. Hence, during the study  
194 there is no perceptual-motor search, only perceptual-motor access.

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

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

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

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

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

### 220 3.1. A rational analysis of the scenarios

221 The rational analysis framework leads us to make a pair of very general predictions con-  
222 cerning the strategies that will be used.

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

227 To work out the specific predictions for the three conditions of our study, Free-Access,  
228 Gray-Box, and Memory-Test, we must first come up with estimates of the efforts of perceptual-

229 motor access versus memory retrieval for each condition. These estimates will enable us to  
230 predict the relative tradeoff of perceptual-motor access versus memory retrieval strategies  
231 between conditions. Based on these effort estimates, we will use our knowledge of cognitive  
232 science in general and memory strength in particular to derive two sets of behavioral predictions  
233 that will be tested in the experiments that follow.

### 234 3.1.1. Access and retrieval efforts

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

245 In terms of perceptual-motor access efforts, both the Memory-Test and Gray-Box groups  
246 have to do the same thing; namely, move visual attention, their eyes, and the cursor to the  
247 correct field in the Information Window, click on its gray box, and perceive a simple word.  
248 Extrapolating from the CPM-GOMS models and data presented by Gray and Boehm-Davis  
249 (2000) for the distances moved here, this time should be between 1,000 and 1,500 ms. This  
250 estimate brackets the estimate for effort obtained from two different modeling approaches. A  
251 CPM-GOMS model developed specifically for this paper (included as Fig. 2) yields an estimate  
252 of 1,280 ms; whereas a modified version of our simple ACT-R model yields an estimate for  
253 effort, *C*, of 1,120 ms. (The modifications simply added a mouse movement and click to the  
254 first ACT-R model.)

255 Effort, measured in time, to retrieve a weak memory is between 500 and 1,000 ms. Although  
256 for important memories, retrieval can be repeatedly cued with the same or different stimulus,  
257 1,000 ms is the upper estimate that Anderson and Lebiere (1998) provide for the time that the  
258 cognitive system will allocate to any one retrieval attempt. This estimate of retrieval time for  
259 weak memories applies to the Free-Access and Gray-Box conditions.

260 Before the Memory-Test group begins a trial they must study and pass a test on the infor-  
261 mation required for that trial. Hence, for them, the memories retrieved are strong. Our estimate  
262 for the time needed to retrieve a strong memory ranges from 100 to 300 ms (Altmann & Gray,  
263 2002; Byrne & Anderson, 2001).

### 264 3.1.2. Strategic predictions by condition

265 The estimates of the effort derived above (and shown in Table 1) enable us to predict the rel-  
266 ative mix of perceptual-motor access or memory retrieval for each condition. The Free-Access  
267 condition should favor perceptual-motor access over memory retrieval. Indeed, the low effort  
268 of perceptual-motor access should rule out use of the slow retrieval of weak and erroneous  
269 memories. However, although we estimated the range of memory retrieval times to be from 500  
270 to 1,000 ms, there may well be some strong memories that can be retrieved faster than this. For

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).

271 the most part, most memories that can be retrieved faster than perceptual-motor access should  
272 be correct. But, because of noise in the memory system some number of erroneous traces may  
273 be relied on.

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

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

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

### 288 3.1.3. Behavioral predictions by condition

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

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

297 For its part, the Free-Access condition should rely on the combination of fast perceptual-  
298 motor access and the fast retrieval of recently encoded, but strong, memories. However, the  
299 500 ms for perceptual-motor access means that, with a boost in activation from random noise,  
300 occasionally a weak and erroneous memory will be quickly retrieved and used in lieu of a  
301 slightly more effortful perceptual-motor access.

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

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

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

### 321 3.2. Auxiliary assumption

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

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

### 334 3.3. Summary of predictions

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

342 perfect knowledge in-the-world, the soft constraints analysis predicts the reliance on imperfect  
343 knowledge in-the-head.

#### 344 **4. Experimental data**

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

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

##### 357 *4.1. Experiment 1*

358 Experiment 1 had three conditions; all used a simulation of a commercial VCR built in  
359 Macintosh Common Lisp™.

##### 360 *4.1.1. Methods*

361 *4.1.1.1. Subjects.* Seventy-two undergraduates, 24 per condition, participated for course cred-  
362 its. Subjects were assigned to conditions randomly in blocks of threes. The experiment took  
363 approximately 30 min. Subjects were individually run.

364 *4.1.1.2. Data collection.* All clicks on any button object in the simulation were time stamped  
365 to the nearest tick (16.67 ms) and saved to a log file along with a complete record of the  
366 information in the VCR's displays (e.g., mode, time, day-of-week, channel, and so on). If  
367 subjects moved the cursor from the VCR window to the Show Information Window the mouse  
368 exit and mouse enter events were recorded and time stamped. Likewise, any clicks on the gray  
369 boxes covering the fields of the Show Information Window were also time stamped and saved  
370 to the log file along with the name of the field and the information it contained. Hence, from  
371 the log file it is possible to reconstruct a complete trace of the subject's interaction with the  
372 VCR and Show Information Window.

373 *4.1.1.3. Procedure.* With minor differences described below, the procedure for all conditions  
374 was the same. The study began with the subject watching as the experimenter programmed the  
375 first trial of show 0. During the show 0 the task-to-device rule hierarchy shown in Fig. 1 was  
376 placed in front of the subject and the experimenter referred to it as s/he programmed the show.  
377 After the first trial the experimenter watched as the subject programmed show 0 to criterion.

378 At that point, the experimenter removed the task-to-device rule hierarchy and left the room  
379 while the subject programmed shows 1–4. (As show 0 was an instruction and practice show, it  
380 is excluded from the analyses reported below.)

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

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

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

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

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

#### 414 4.1.2. Results and discussion

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

418 4.1.2.1. Performance measures: errors. Trials-to-criterion. A trial started when the subjected  
419 clicked the START TRIAL button and continued until the STOP TRIAL button was clicked.

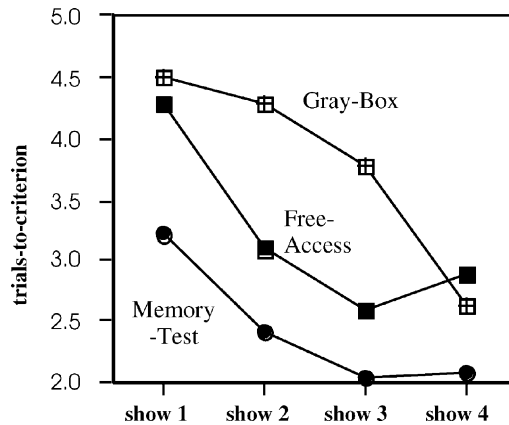


Fig. 4. Trials-to-criterion for experiment 1. Subjects were required to program each show to the criterion of two successive correct trials. Hence, for shows 3 and 4 the Memory-Test group is close to the minimum number of trials possible.

420 Trials for each show continued until the show was programmed correctly for two successive  
 421 trials. Given that, in each condition, show information was readily available in-the-world, we  
 422 might have expected all subjects in all conditions to have spent a maximum of two trials per  
 423 show. A trials-to-criterion score of greater than two, reflects the number of trials that, when the  
 424 subject clicked the STOP TRIAL button, were not correctly programmed. Hence, we interpret  
 425 trials-to-criterion greater than two as reflecting a reliance on imperfect memory in lieu of  
 426 accessing knowledge in-the-world.

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

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

438 *Goal suspension.* The trials-to-criterion measure focused our attention on the number of trials  
 439 that ended in error; that is, the number of trials that ended with a show being incorrectly pro-  
 440 grammed. The more shows that were incorrectly programmed the greater the trials-to-criterion.  
 441 In contrast, for goal suspension we examine errors that were made, but latter detected and cor-  
 442 rected on trials that ended successfully.

443 The measure of goal suspension is derived from Gray's (2000) goal-structure analysis of  
 444 errors of performance<sup>2</sup>. For the VCR simulation there are eight fields that must be set to  
 445 correctly program the VCR; day-of-week, channel, start-hr, start-10 min, start-min, end-hr,  
 446 end-10 min, end-min. Given the structure of the device, the measure of goal suspension is quite

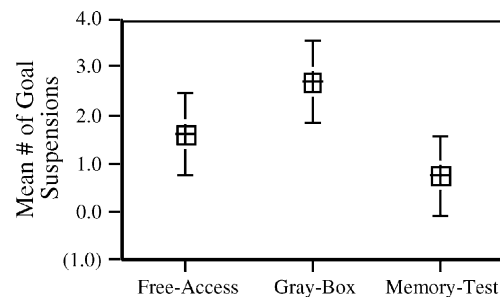


Fig. 5. Mean goal suspensions per subject across the three conditions. statistical significance bars (SSBs) show the pairwise statistical significance between means.

447 simple: Once a subject starts to change a setting, how often was it abandoned before being  
 448 correctly completed? For example, if for show 2 the to-be-set channel was 21, and the current  
 449 channel was 11, then if the subject began setting the channel but stopped before 21 (e.g., going  
 450 off to set the day-of-week), then this is one goal suspension.

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

462 Goal suspensions are a rarity. Examining patterns of goal suspensions requires that a vast  
 463 quantity of correct data be collected and parsed. Across all three conditions of experiment 1,  
 464 36,877 mouse clicks were collected and time stamped on correct trials. These mouse clicks  
 465 were parsed into 12,560 goals using the action-protocol analyzer developed by Fu (2001) with  
 466 the task-to-device rule hierarchy shown in Fig. 1. Less than 1% of all goals, 122, resulted in  
 467 goal suspensions. For each group the mean number of goal suspensions per subject is shown  
 468 in Fig. 5.

469 An overall ANOVA produced a marginally significant effect,  $F(2, 69) = 2.64, p = .078$ .  
 470 The statistical significance bars (SSBs) in Fig. 5 are based on planned comparisons. If two SSBs  
 471 look different (i.e., they do not overlap), the corresponding pairwise comparison is different  
 472 (at the .05 level of significance adopted for this study) (for more information on SSBs see  
 473 Schunn, 2000). As indicated by Fig. 5, the Gray-Box condition made significantly more goal  
 474 suspensions than did the Memory-Test condition, but there were no significant differences  
 475 between the other comparisons.

476 A  $\chi^2$  comparison that looked at whether or not each subject made goal suspensions, was  
 477 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.

478 4.1.2.2. *Discussion of performance measures.* The first two dependent measures, trials-to-  
479 criterion and goal suspensions, yield a consistent pattern. The Memory-Test condition is best,  
480 and the Gray-Box is worst with the Free-Access condition somewhere in the middle. These  
481 data present us with an interesting quandary. All groups had access to all show informa-  
482 tion at all times, yet they made errors that kept them in the study longer than they needed  
483 to be.

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

494 4.1.2.3. *Process measure: accesses of knowledge in-the-world.* Our process measure counts  
495 the number and the pattern of information accesses to the Show Information Window. For the  
496 Memory-Test and Gray-Box conditions, each click on a gray box was counted. The pattern of  
497 when information was accessed versus when the information was programmed was derived  
498 from the log files.

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

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

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

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

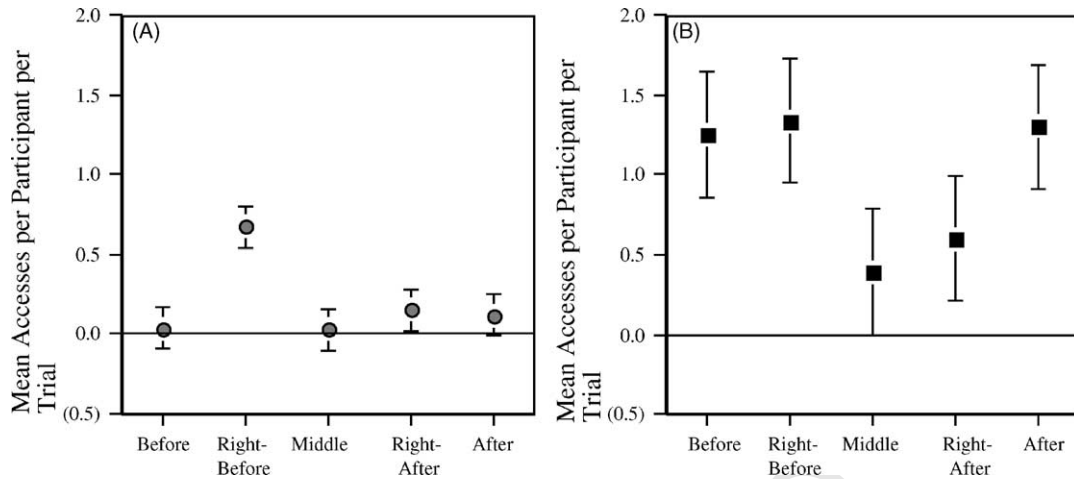


Fig. 6. (A) For the Gray-Box condition from experiment 1, the graph shows the mean accesses per subject per trial, SSBs, based on 24 subjects, show the pairwise statistical significance between means. (B) For the Free-Access condition from experiment 2, the graph shows the mean accesses per subject per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

521 act was to access the channel information, this access was classified as *right-after*. Any later  
 522 access of an information field was classified as *after*.

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

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

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

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

## 545 4.2. Experiment 2

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

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

### 552 4.2.1. Method

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

557 4.2.1.1. *Subjects.* We report results from the first eight undergraduates who we could success-  
558 fully eye track. All subjects, whether or not they could be eye-tracked, received course credit  
559 for their participation. Because of the necessity to calibrate the eye-tracker on each subject,  
560 experiment 2 took approximately 45 min.

561 4.2.1.2. *Eye tracking.* Eye tracking was performed using an ASL 504 remote optics eye tracker.  
562 Head movements were tracked using a Flock-of-Birds<sup>TM</sup> magnetic head tracker. Eye data was  
563 sampled and saved to a log file 60 times per second (once every 16.67 ms).

564 Fixations were determined using the algorithm developed by [Karsh and Breitenbach](#)  
565 (1983). Basically, we say that a fixation occurs when at least six consecutive data points  
566 fall within a  $3 \times 4$  pixel rectangle (where the definition of “consecutive” points is that they  
567 have to be less than 32 ms apart). Areas of interest were created around each information  
568 box. Consecutive fixations in the same area of interest were counted as a single  
569 access.

### 570 4.2.2. Results

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

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

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

584 significantly differs from the middle category and is marginally different from the right-after  
585 category. There are no differences between the middle and right-after categories.

#### 586 4.3. Discussion of information access in experiments 1 and 2

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

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

### 604 5. General discussion

#### 605 5.1. Success of predictions—support for hypotheses and assumptions

606 Plugging Table 1 estimates of perceptual-motor and cognitive effort into our rational analysis  
607 led us to derive soft constraints in different experimental conditions and make several predic-  
608 tions concerning performance and process. For performance we predicted that the Memory-Test  
609 condition would be best and Gray-Box the worst with the Free-Access group somewhere in  
610 the middle. We found this rank-order with both performance measures. For trials-to-criterion  
611 there were significant differences between Memory-Test and each of the other two conditions,  
612 but not significant differences between these two. For goal suspensions planned comparisons  
613 revealed a significant difference between Memory-Test and Gray-Box with no other significant  
614 between group differences.

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

620 The process measure directly supports the prediction that on many occasions when the  
621 Gray-Box or Free-Access condition might have been expected to rely on knowledge in-the-  
622 world they relied, instead, on knowledge in-the-head. For example, while programming (mid-

623 dle) a setting and right-after programming a setting the Gray-Box condition made almost no  
624 accesses of the Show Information Window (see Fig. 6A). The Free-Access condition made  
625 only slightly more middle and right-after accesses, averaging approximately 1/2 an access per  
626 person per trial (see Fig. 6B).

627 These behavioral measures support our predictions based on our notion of least effort trade-  
628 offs as soft constraints that govern non-deliberate selection of microstrategies in interactive be-  
629 havior. The success of our analysis in predicting experimental outcomes provides indirect sup-  
630 port for our auxiliary assumption. For patterns of interactive behavior lasting between 500 and  
631 1,000 ms, time is a reasonable surrogate measure for cognitive as well as for perceptual-motor  
632 effort. Unless our estimates of the duration of memory and perceptual-motor processes are  
633 way off, the success of our predictions suggests that the time spent retrieving something from  
634 memory is weighed the same as time spent in perceptual-motor activity.

## 635 5.2. Implications

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

647 It may not be completely surprising that such small differences in perceptual-motor efforts  
648 affected strategy (Wickens, 1992), but we are surprised that they influenced performance.  
649 Indeed, the most striking aspect of the between-group differences in performance is that all  
650 were avoidable. All performance differences can be traced to differences in willingness to either  
651 memorize or access show information. For each trial the Memory-Test group had quick and  
652 reliable access to show information in memory. The other groups made more errors that resulted  
653 in more trials-to-criterion and more goal-suspensions. Apparently, verification is lower effort—  
654 and hence more likely—if based on knowledge in-the-head rather than accessing knowledge  
655 in-the-world.

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

### 661 5.2.1. Implications for routine interactive behavior

662 Although the level of perceptual-motor and memory effort manipulated in these studies  
663 was representative of that encountered in many human-system interactions, this effort is much

664 lower than that involved in many others. For example, for a typical process control operator,  
665 accessing knowledge in-the-world may require more than simply clicking on a gray box, it may  
666 require getting out of a seat and moving across the room to an information display. Similarly,  
667 unlike the situations we studied, visually busy web pages impose a substantial search effort  
668 on accessing knowledge in-the-world. Hence, for much routine interactive behavior, we would  
669 expect the effort-benefit tradeoff to favor accessing imperfect knowledge in-the-head over  
670 perfect knowledge in-the-world.

### 671 5.2.2. *Implications for design*

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

### 686 5.2.3. *Implications for embodied cognition*

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

692 We have heard and agree to all of these points. Indeed, our point is precisely that the most  
693 adaptive pattern of interactive behavior is one that is the least effort given the current task and  
694 the current task environment. Change the environment or the task then another pattern will be  
695 most adaptive. However, like Scaife and Rogers (1996), rather than appealing to hand-waving  
696 theories or intuitive accounts of what should be easier for people to do, we call for an account  
697 of the control of interactive behavior; that is, the ways in which central cognition orchestrates  
698 processes such as mouse movements, eye movements, perception, shifts in attention, memory  
699 encoding, and memory retrieval.

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

704 We argue for a more nuanced approach, one that does not presume the privileged status of  
705 any location or type of operation. Indeed, we would rephrase Wilson's third sense of embodied

706 cognition to read, “the cognitive control of interactive behavior minimizes effort by using a  
707 least effort combination of the mechanisms available to it.” All mechanisms or subsystems are  
708 on the table. There is no reason to think that one mechanism or subsystem has a privileged  
709 status in relation to another.

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

## 713 6. Conclusions: soft constraints in interactive behavior

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

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

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

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

## 733 Notes

- 734 1. A compressed file (StuffIt™ format) containing two Lisp files and an Excel work-  
735 book may be downloaded from the annex maintained by the *Cognitive Science Society*,  
736 see <http://cogsci.psy.utexas.edu/supplements/>. The Lisp files are written in Macintosh  
737 Common Lisp (MCL) but use the uniform-interface standard for ACT-R and hence the  
738 windows created as well as the models should run under ACL Lisp (though this has not  
739 been tested). Each Lisp file implements one of the models discussed in this section. The  
740 Excel workbook includes sample runs of each model and estimates of effort after 1, 10,  
741 100, and 1,000 cycles.
- 742 2. 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*.

743 **Acknowledgments**

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

747 **References**

- 748 Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interfer-  
 749 ence. *Psychological Science*, *13*(1), 27–33.
- 750 Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- 751 Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences*, *14*(3), 471–517.
- 752 Anderson, J. R., Bothell, D., Byrne, M. D., & Lebiere, C. (in press). An integrated theory of the mind. *Psychological*  
 753 *Review*.
- 754 Anderson, J. R., & Lebiere, C. (Eds.). (1998). *Atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum  
 755 Associates.
- 756 Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*,  
 757 396–408.
- 758 Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive*  
 759 *Neuroscience*, *7*(1), 66–80.
- 760 Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition.  
 761 *Behavioral and Brain Sciences*, *20*(4), 723–742.
- 762 Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect  
 763 time-sharing. *Psychological Review*, *108*(4), 847–869.
- 764 Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human–computer interaction*. Hillsdale, NJ:  
 765 Lawrence Erlbaum Associates.
- 766 Cary, M., & Carlson, R. A. (1999). External support and the development of problem-solving routines. *Journal of*  
 767 *Experimental Psychology-Learning Memory and Cognition*, *25*(4), 1053–1070.
- 768 Cary, M., & Carlson, R. A. (2001). Distributing working memory resources during problem solving. *Journal of*  
 769 *Experimental Psychology-Learning Memory and Cognition*, *27*(3), 836–848.
- 770 Ehret, B. D. (2000). Learning where to look: The acquisition of location knowledge in display-based interaction.  
 771 *Dissertation Abstracts International: Section B: The Sciences & Engineering*, *60*(10-B), 5239.
- 772 Ehret, B. D. (2002). Learning where to look: Location learning in graphical user interfaces. *CHI Letters*, *4*(1),  
 773 211–218.
- 774 Frohlich, D. M. (1997). Direct manipulation and other lessons. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.),  
 775 *Handbook of human–computer interaction* (2nd ed., pp. 463–488). New York: Elsevier.
- 776 Fu, W.-T. (2001). ACT-PRO: Action protocol tracer—a tool for analyzing discrete action protocols. *Behavior*  
 777 *Research Methods, Instruments, & Computers*, *33*(2), 149–158.
- 778 Gray, W. D. (2000). The nature and processing of errors in interactive behavior. *Cognitive Science*, *24*(2), 205–248.
- 779 Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use  
 780 in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied*, *6*(4), 322–335.
- 781 Gray, W. D., & Fu, W.-T. (2001). Ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head:  
 782 Implications of rational analysis for interface design. *CHI Letters*, *3*(1), 112–119.
- 783 Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting  
 784 and explaining real-world performance. *Human–Computer Interaction*, *8*(3), 237–309.
- 785 Gray, W. D., & Salzman, M. C. (1998). Damaged merchandise? A review of experiments that compare usability  
 786 evaluation methods. *Human–Computer Interaction*, *13*(3), 203–261.
- 787 Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human–Computer Inter-*  
 788 *action*, *1*(4), 311–338.

- 789 Karsh, R., & Breitenbach, F. W. (1983). Looking at looking: The amorphous fixation measure. In R. Groner, C.  
790 Menz, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and psychological functions: International views*  
791 (pp. 53–64). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 792 Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with  
793 application to human–computer interaction. *Human–Computer Interaction*, 12(4), 391–438.
- 794 Kieras, D. E., & Meyer, D. E. (2000). The role of cognitive task analysis in the application of predictive models  
795 of human performance. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive task analysis*  
796 (pp. 237–260). Mahwah, NJ: Lawrence Erlbaum Associates.
- 797 Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*,  
798 11, 65–99.
- 799 Neth, H., & Payne, S. J. (2001). Addition as interactive problem solving. In J. D. Moore & K. Stenning (Eds.),  
800 *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 698–703). Hillsdale,  
801 NJ: Lawrence Erlbaum Associates.
- 802 Norman, D. A. (1989). *The design of everyday things*. New York: DoubleDay.
- 803 O’Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving  
804 and learning. *Cognitive Psychology*, 35, 34–70.
- 805 O’Hara, K. P., & Payne, S. J. (1999). Planning and the user interface: The effects of lockout time and error recovery  
806 cost. *International Journal of Human–Computer Studies*, 50(1), 41–59.
- 807 Oaksford, M., & Chater, N. (Eds.). (1998). *Rational models of cognition*. New York: Oxford University Press.
- 808 Payne, S. J., Howes, A., & Reader, W. R. (2001). Adaptively distributing cognition: A decision-making perspective  
809 on human–computer interaction. *Behaviour & Information Technology*, 20(5), 339–346.
- 810 Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal*  
811 *of Human–Computer Studies*, 45(2), 185–213.
- 812 Schunn, C. D. (2000). Statistical significance bars (SSB): A way to make graphs more interpretable. Manuscript  
813 submitted for publication.
- 814 Schweickert, R., Fisher, D. L., & Proctor, R. W. (2003). Steps toward building mathematical and computer models  
815 from cognitive task analyses. *Human Factors*, 45(1), 77–103.
- 816 Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cog-*  
817 *nitive Science*, 26(4), 469–501.
- 818 Shneiderman, B. (1982). The future of interactive systems and the emergence of direct manipulation. *Behaviour*  
819 *and Information Technology*, 1(3), 237–256.
- 820 Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.
- 821 Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.
- 822 Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636.