Title
Memory versus Perceptual-Motor Tradeoffs in a Blocks World Task

Permalink
https://escholarship.org/uc/item/5mn4s2zt

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

ISSN
1069-7977

Authors
Fu, Wai-Tat
Gray, Wayne D.

Publication Date
2000

Peer reviewed

Memory versus Perceptual-Motor Tradeoffs in a Blocks World Task

Wai-Tat Fu (wfu@gmu.edu)
Wayne D. Gray (gray@gmu.edu)
Human Factors & Applied Cognition
George Mason University
Fairfax, VA 22030 USA

Abstract

Using information *in-the-world* as *external memory* may be a low-cost alternative to internal memory: storage is free, and retrieval is often quick (involving a saccade) and reliable. However, when the cost of accessing external information increases, *in-the-head* storage and retrieval may become the least-cost solution. We employ the rational analysis framework (Anderson, 1990) to study the effect of varying the cost of information access on interactive behavior. Increasing the cost of information access induced a switch from information in-the-world (the perceptual-motor strategy) to information in-the-head (the memory strategy). Given the effort and unreliability of internal storage, the threshold for switching from an in-the-world to an in-the-head strategy is surprisingly low.

Introduction

Information stored *in-the-world* can be considered as *external memory* (O'Regan, 1992). Information is retrieved from external memory via visual perception as rendered by the appropriate saccades and fixations. Recent research has suggested that when information in-the-world is readily accessible, internal storage is not needed (Ballard, Hayhoe, & Pelz, 1995); perceptual-motor strategies will be deployed to reacquire information as needed. However, when the cost of information access was increased from a simple saccade to a head movement, the perceptual-motor strategy was replaced with a strategy that placed task-relevant information into working memory (Ballard et. al., 1995). This suggests that the decision to store information in-the-head versus in-the-world is sensitive to least-cost considerations.

The rational analysis framework

One explanation for this kind of trade-off was given by Anderson (1990). Anderson casts human memory as an optimization process. In his rational analysis framework, the goal of human memory is to retrieve knowledge that would allow us to perform the task we are currently facing. The optimization process maximizes the *expected utility* of the memory system by balancing the cost of memory search against an assumed constant expected gain[1] of retrieving a desired memory item for the current task. A clear cost of memory search is time (and possibly a metabolic cost associated with time). Under Anderson's rational analysis framework, the human memory system would search a memory structure until the probability of getting the desired memory item (the expected gain) is lower than the cost of further search (i.e., when the expected utility becomes negative).

If information in the external environment can be considered as an external memory store, the cost in searching for the relevant information in the external environment can be taken as the "memory" search cost as in Anderson's rational analysis framework. In most tasks, the information stored in the external environment is continuously available (high expected gain and fixed expected cost).

If the only cost associated with internal memory were a search cost, then we would expect that in most situations internal search would be faster than external search. However, for internal memory a significant additional cost is internal storage (encoding). Storage costs would seem to be particularly problematic in the type of real-time, dynamic tasks studied by Ballard and associates. For example, in a task that required frequent memory updates, Altmann and Gray (2000) estimated that approximately 10 cycles of encoding with a duration of approximately 100 msec per cycle are needed to encode an item so that it can be retrieved 5,000 msec later. In contrast, the time for a saccade and dwell is typically estimated as 230 msec (Card, Moran, & Newell, 1983, pp. 25-28).

Compared to a memory strategy that includes encoding plus retrieval, a saccadic eye movement to a known location has a much lower time cost. Therefore, under many conditions, the expected utility of using the external environment as external memory is much higher than that of the internal human memory system. However, when the cost of information access from the external environment is high enough, the expected utility of external memory would be lower than that of internal memory. In this case, the rational analysis framework would predict a shift from external memory to internal memory. In other words, people would be more likely to adopt a memory strategy than a perceptual-motor strategy.

Unlike retrieving an item from a known external location with a saccade and dwell, retrieving an item from memory is

[1] The expected gain is defined as the product of P and G, where P is the estimated probability that the target memory item can be found, and G is the gain associated with retrieving the target memory item. If C is the memory search cost, then expected utility E = PG - C.

subject to interference from previously encoded as well as other currently encoded items. If we make the additional assumption that the strength of an encoded trace fluctuates as a function of noise (Altmann & Gray, 1999; Anderson & Lebiére, 1998), then retrieval from memory may take longer and is more error prone than the corresponding retrieval from the external environment. The rational analysis framework suggests that searching for an item should stop as soon as the expected gain from finding the item is less than the cost of searching. Therefore given an assumed constant expected gain, the higher the search cost of a memory item, the fewer items would be searched for and inspected before the memory system would stop searching. Since the more items the memory system considers, the more likely that the target item can be found (thus improving accuracy), increasing the search cost implies a decrease in accuracy; that is, an increase in errors. Therefore the rational analysis framework not only predicts that increasing the cost of information access in the external environment would induce a shift from external memory to internal memory, but also an increase in errors.

In this paper, we employ the rational analysis framework to study the effect of varying the cost of information access on interactive behavior. Specifically we test two predictions that we have derived from the rational analysis framework: that an increase in the perceptual-motor cost of information access will induce a shift from an external to an internal memory strategy, and that this switch will occur even though the internal search is difficult and error prone.

Experiment

The blocks world task is based on the paradigm used by Ballard et al., (1995). The task is to copy a pattern of colored blocks shown in the *target* window to the *workspace* window, using the colored blocks in the *resource* window (for our version see Figure 1).

To do the task, participants have to remember three pieces of information: (a) the color of the block to be copied, (b) the position of the block to be copied, and (c) which blocks have or have not been copied. The first two pieces have to be obtained from the target window whereas the third piece has to be obtained by comparing the target window with the workspace window.

Ballard reported a point-of-gaze (POG) sequence of target window, resource window, target window, workspace window (TRTW). The implication of this sequence is that during the first POG to the target window (T) subjects encoded the color of the block and then picked up a block from the resource window (R). On the next POG to the target window (T) subjects encoded the block's location in the pattern. They then moved to the workspace window (W) and placed the block in the appropriate location.

As the effort needed to acquire information from a window increased from a POG to a head movement the sequence tended to change to TRW. In this case, the implication is that subjects encode both the color and the position during the first (and only) POG to the target window (T). The encoded trace persists as the subject acquires the block from the resource window (R) and places it in the workspace window (W).

Unlike Ballard et al., in our Block World task all three windows were covered by gray boxes. Throughout the task only one of the windows could be uncovered at a time. The resource and workspace windows were uncovered by moving the mouse cursor into the window. They were covered again when the mouse cursor left the window. The effort required to uncover the target window varied between each of our three conditions.

To access the information in the target window participants could adopt either a predominately perceptual-motor or a predominately memory strategy. As per Ballard et al.'s TRTW strategy, the predominately perceptual-motor strategy would entail one uncovering at the target window to obtain color information and another to obtain position information. In contrast, a predominately memory strategy (TRW) would entail one uncovering at the target window to obtain both color and position. Deciding which blocks remained to be copied would entail a second set of strategies. On these strategies Ballard is silent. However, a predominately perceptual-motor strategy might entail multiple quick uncoverings between the target and workspace window. A predominately memory strategy might entail encoding the color and position of multiple blocks at one glance.

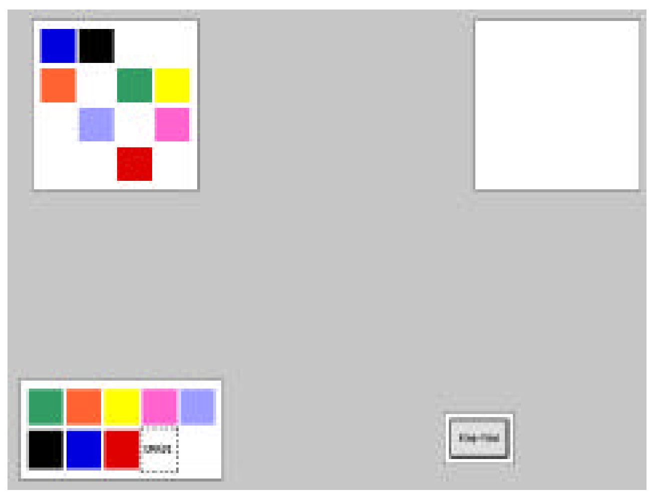

Figure 1. The blocks world task. In the actual task all windows are covered by gray boxes and at any time only one window can be uncovered. The window at the top left is the target window, at the bottom the resource window, and at the top right the workspace window.

Method

Participants

Forty-eight George Mason University undergraduates participated in the study for course credit and were randomly assigned to one of the three experimental conditions.

Equipment and software

The experiment was written in Macintosh Common Lisp and was conducted with a Macintosh PowerPC connected to an extended keyboard, a mouse, and a 17-inch monitor. All mouse movements and keypresses were recorded and saved to a log file with 16.67 msec accuracy.

The blocks (48 x 48 pixels) that constitute each pattern were randomly chosen with the constraint that no color was used in one pattern more than twice. The blocks were placed at random in the target window's 4 x 4 grid. The workspace window was the same size as the target window and contained the same, non-visible, 4 x 4 grid.

Design and Procedure

The three conditions were designed to vary the cost of uncovering the target window. In the *low-cost* condition, participants had to press and hold down a function key. (Participants were asked to use different hands for the keyboard and the mouse.) The target window remained uncovered until they released the key, or until the mouse cursor entered either the workspace or resource window. Once the target window was covered, to uncover it again participants had to release the key and press it again (this is to avoid the strategy of holding the key down throughout the task). In the *control* condition, the conditions for uncovering the target window were the same as for the workspace and resource windows. The target window was uncovered when the mouse cursor entered the window. The *high-cost* condition was similar to the control condition, except that participants had to move the mouse cursor inside the target window and endure a one second lockout before the target window was uncovered.

To select a block, participants moved the mouse cursor to the resource window and mouse clicked on the desired colored block. The mouse cursor then changed to a small version (16 x 16 pixels) of the selected block (eliminating the need to remember its color). To place a block in the workspace window, the cursor was moved to that window (which was then uncovered), moved to the desired position, and then clicked. When the participants believed that the pattern had been copied to the workspace window, they press the "Stop-Trial" button. A feedback window indicated whether the copied pattern matched the target pattern. If the pattern was different, participants were required to go back to finish the task before they could move on to the next pattern.

At the beginning of the experiment, instructions were given and participants were led through one demonstration trial by the experimenter. Participants then completed 40 trials. The whole experiment lasted about 45 minutes.

Results

Trial Time

The first ten trials of the experiment were considered practice and were excluded from further analyses. Analysis of variance (ANOVA) of time for condition by trial showed significant main effects of condition (F (2, 45) = 9.11, p = 0.0005, MSE = 726), and of trial (F (29, 1305) = 131.6, p < .0001, MSE = 53.2). There was no interaction between trial and condition. To determine whether the main effect of conditions was solely due to the one second delay in the high-cost condition, the per trial time in this condition was adjusted by subtracting the amount of delay for each time the target window was uncovered. After the adjustment, the main effect of condition was not significant (F (2, 45) = .969, p = .39, MSE = 564). However, the main effect of trial remained significant (F (29, 1305) = 120.1, p < .001, MSE = 46.0). Orthogonal linear contrasts showed a significant linear downward trend of time across trials for the low-cost condition (p = .0001), control condition (p = .0001), and high-cost condition (p = .0001), suggesting speed-up across trials. No other higher order trends were significant. The interaction between trials and conditions was not significant.

Use of the Target Window

The trial time results seem to suggest no difference between conditions. However, detailed analyses revealed the effects of the cost of information access. An ANOVA on the number of times the target was uncovered showed a significant main effect of condition (F (2, 45) = 10.17, p = .0002, MSE = 159). Planned comparisons revealed a significant difference between the high-cost and control (p = .0045), as well as high-cost and low-cost conditions (p < .0001) (See Figure 2). Subjects in the high-cost condition uncovered the target window significantly fewer times than the other two conditions. However, there was no significant difference between the low-cost and control condition.

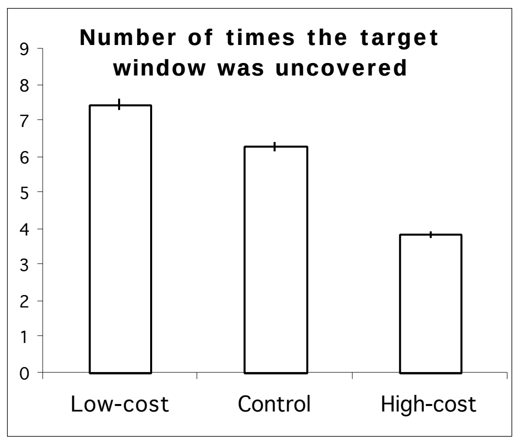

Figure 2. Mean number of times per trial that subjects uncovered the target window.

ANOVA on the time subjects spent looking at the model showed that there were significant main effects of conditions (F (2, 45) = 20.6, p < .0001, MSE = 300), with the high-cost condition significantly spending more time than the low-cost condition (p < .0001), or the control condition (p < .0001). The difference between the low-cost and control condition was not significant. Overall, there was a significant downward linear trend of time spent on the target window across trials (p = .0002). However, orthogonal linear contrasts showed that the downward linear trends for the high-cost (p = .02) and control condition (p = .001) were significant, but that the trend for low-cost condition (p = .10) was not (see Figure 3).

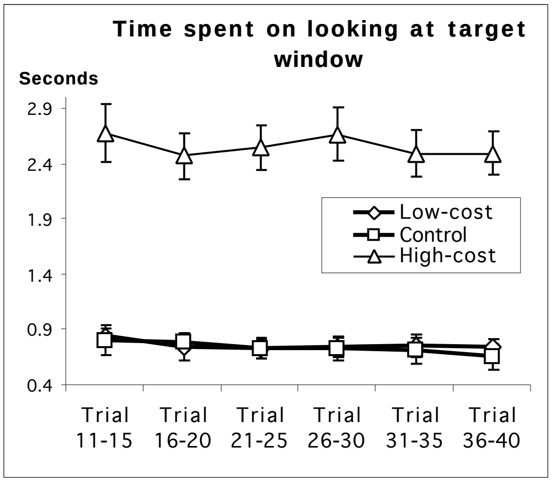

Figure 3. Time spent on looking at the target window.

For each trial, we looked at how many colored blocks the subjects copied following each of their first four accesses of the target window. We conducted a 3 x (4 x 30) ANOVA on conditions, the nth (1 to 4) uncovering of the target window, and trial (11-40). There were significant main effects of conditions (F (2, 45) = 19.5, $p < .0001$, MSE = 7.75), and the nth uncovering of the window (F (3, 135) = 39.5, $p < .0001$, MSE = 13.3). The interaction between conditions and the nth window uncovering was significant (F (6, 135) = 7.8, $p < .0001$, MSE = 13.2) (see Figure 4). The main effect of trials was not significant. No other interactions were significant.

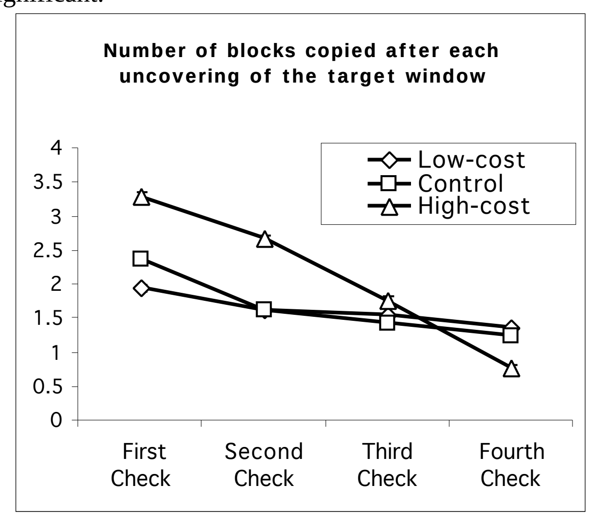

Figure 4. The number of colored blocks copied after each uncovering of the target window.

Summary. The analysis of how the target window was used suggest that although there was no significant difference in trial times between conditions, the strategies used by the subjects in the high-cost condition were very different from the other two conditions. Subjects in the high-cost condition uncovered the target window a fewer number of times and copied more colored blocks after each uncovering at the target window. Figure 4 illustrates the interaction between condition and the number of blocks copied in the first four uncoverings of the target window. Clearly, the significant decrease in the number of blocks copied within a trial is at least partially due to the decreasing number that remained to be copied. However, this potential artifact does not explain the significant interaction. Subjects in the high-cost condition tended to copy more blocks in the first few accesses at the target window. In contrast, in the low-cost condition, subjects tended to copy the same number of blocks after each access. This suggests that subjects in the high-cost condition tended to adopt a memory strategy - memorizing the information of the colored blocks to reduce their reliance on the external display. On the other hand, as the cost of information access was low subjects in the low-cost condition relied more on a perceptual-motor strategy - getting the information from the display when needed.

With practice, subjects in the high-cost condition spent less time looking at the target window, but copied more colored blocks after each uncovering. This increase in efficiency (measured in terms of the time required to memorize the information of a fixed number of colored blocks) partly explains the speed-up across trials in the high-cost condition. However, the same trend was not observed in the low-cost condition. It seems that practice had no significant effect on the use of the target window in the low-cost condition (in terms of number of times they uncovered at the target window, number of colored blocks copied after each uncovering, and time spent on the target window per uncovering).

Strategies

A finer-grained analysis is needed to understand the actual differences in strategies used. For each copied block, we extracted the sequence in which subjects uncovered windows. In the notation below, these sequences are abbreviated using the first initial of the window name. Wb indicates that they uncovered the workspace window and placed a block, Wu indicates that they uncovered the workspace window but did not place a block. For example, TRWb refers to a target-source-workspace windows sequence that ends with the placement of a block.

Table 1. Strategies used by subjects. T = uncover target window, R = uncover and pick a colored block from resource window, Wb = put selected colored block to workspace window, Wu = uncover workspace window. For example, TWuRWb represents the strategy in which the subject uncovered the target window, uncovered the workspace window, went to the resource window, picked up a colored block, and put the colored block in the workspace window.

Low-cost		Control		High-cost	
Strategy		Strategy		Strategy	
TRWb	53%	TRWb	49%	TRWb	38%
RWb	33%	RWb	38%	RWb	58%
TRTWb	5%	TRTWb	7%	TWuRWb	1%
Total	5537		5496		5752

Table 1 shows the three most common sequences used by the subjects in the three conditions (as well as the percentage of the total that these sequences represent). We can see that the top 2 sequences (TRWb and RWb) constituted almost 90% of all the sequences used[2].

Although the two dominant strategies were the same, the effect of information access cost on strategy used was clearly seen. With increasing cost, the use of the TRWb strategy decreased, while the use of RWb increased. This change of strategy nicely indicates the shift of reliance from external to internal memory. This is consistent with the results on the use of target window described above. With increasing cost of information access, subjects tended to uncovered the target window less, spent more time per uncovering (time that we presume was spent encoding more information into internal memory), and used the pure memory strategy (RWb) more.

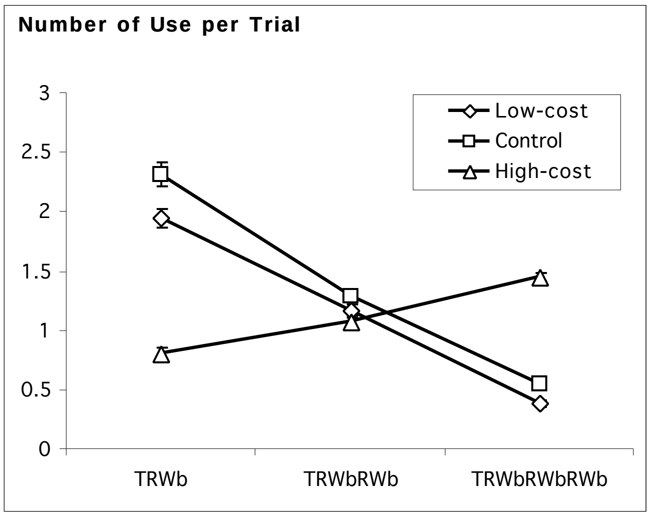

Figure 5. Number of use of strategies per trial. T = uncovering the target window, R = uncovering the resource window and picked a colored block, Wb = putting the colored block in the workspace window.

To further understand the strategy change across trials, we looked at three combinations of the two dominant strategies. We extracted the number of times subjects used either the TRWb, TRWbRWb, or TRWbRWbRWb strategy[3] in each of the trials. This analysis is similar to that shown in Figure 4. It captures how often subjects placed 1, 2, or 3 blocks following a single glance at the target window. A 3 x (3 x 30) ANOVA on conditions, strategies, and trials found no significant main effect of condition (F (2 ,45) = 1.92, p = .16, MSE = 16.0), nor trials (F (29, 1305) =. 835, p < .72, MSE = 24.2). However, the main effect of sequence was significant (F (2, 90) = 14.6, p < .0001, MSE = 19.8), as was the interaction between sequence and condition (F (4, 90) = 10.9, p < .0001) (see Figure 5). No other interactions were significant. Planned comparisons showed that all differences between the three sequences were significant (p < .05).

The analysis further confirms the shift from the use of external to internal memory with increasing information access cost. The TRWb strategy, in contrast to the TRWbRWb and TRWbRWbRWb strategies, allowed subjects to acquire only the information necessary to copy the next colored block from the target window (external memory). Figure 5 shows that in the high-cost condition, subjects used the TRWbRWbRWb significantly more than the TRWb strategy. It suggests that subjects tended to transfer more information from the target window (external memory) to internal memory, and performed the task based on the information retained in internal memory.

Errors and Comparisons

To test the predictions of errors from our rational analysis, we looked at the way in which different strategies affected the patterns of errors made by subjects. An error could involve placing a block with the wrong color, placing a block in the wrong position, or both. Only errors made before the subjects clicked on the "Stop-Trial" button were counted. A 3 x 30 ANOVA of the number of errors made per trial on conditions and trials was conducted. There was a significant main effect for conditions (F (2, 45) = 11.6, p <.0001, MSE = 1.39), but not for trials (F (29, 1305) = 1.1, p = . 33, MSE =. 75). The interaction between conditions and trials was not significant (F (58, 1305) = 1.2, p = .10). The row labeled “Errors” in Table 2 shows the mean number of errors made by the subjects. The low error rates were not surprising given the simplicity of the task. However, there were significant differences between conditions. Planned comparisons showed that subjects in the high-cost condition made significantly more errors than the other two conditions (p < .001 for low versus high, p < .0001 for control versus high). This result supports our second prediction: that increase in information access cost increases errors.

Table 2. Mean number of errors, errors uncorrected, and comparison episodes per trial before the subjects thought they were done (when they pressed the "Stop-Trial" button.)

	Conditions (cell means)		
Dependent Variable	Low-cost	Control	High-cost
Errors	.41	.33	.68
Uncorrected	.10	.08	.14
Comparisons	.57	.31	.17

To find out whether there were differences in the subjects' ability to detect errors, we conducted a 3 x 30 ANOVA on the number of uncorrected errors after the subjects clicked on the "Stop-Trial" button by condition and trial. As suggested by the middle row of Table 2, there was no significant main effect on conditions (F (2, 45) = 1.26, p = .29, MSE = .365). However, there was a significant main effects of trial (F (29, 1305) = 1.54, p = .03, MSE = .173). There was a significant downward linear trend on the number of uncorrected errors across trials (p = .006).

[2] Interestingly, TRTW, the dominant strategy described in Ballard et al (1995), was not one of the dominant strategy in this task. The difference might be that all windows in this task were covered by gray boxes, and the lower-cost saccadic strategy described in Ballard et al (1995) was not supported.

[3] In our categorization, these categories were mutually exclusive. Therefore, the run TRWRWRW was categorized as an instance of the TRWRWRW strategy, but was not included in the count for TRWRW or TRW.

The decrease of the number of uncorrected errors (without any significant increase in the number of errors made) suggested that with practice, subjects in the high-cost condition became better at detecting and correcting their errors. We therefore turned our focus on how often the subjects compared the pattern in the workspace with that in the target window. The comparisons between the two windows not only served the function of error detection, but could also let the subjects keep track of what blocks had or had not been copied. As described before, this information was another critical piece of information that had to be remembered to do the task.

The number of comparison episodes was extracted. A comparison episode started when the participant went from the workspace to the target window (or vice versa) without having a block selected. Any consecutive uncoverings of the workspace and target window were counted as part of the same comparison episode. An ANOVA of the number of comparison episodes showed a significant main effect on conditions ($F(2, 45) = 6.3$, $p = .004$, $MSE = 3.17$), with the low-cost condition having significantly more comparison episodes than the other two conditions (see the bottom row of Table 2). The result again confirmed our prediction: when the cost of information access is low, people prefer to adopt a perceptual-motor strategy to memory strategy.

To summarize, the memory strategy adopted by subjects in the high-cost condition seemed to be more error-prone than the perceptual strategy in the low-cost condition. However, in all conditions most errors were detected and corrected. There was therefore no significant difference in the number of uncorrected errors between conditions. The differences in number of comparison episodes revealed another aspect of the strategy difference. In the low-cost condition, subjects made many more comparisons of the workspace and the target window, further supporting the hypothesis that they did not keep track of which blocks had been copied. In contrast, the number of comparison episodes in the high-cost condition was much lower, suggesting that subjects stored the information in working memory, reducing the reliance on the external environment. A second function served by these comparisons might have been to detect and correct errors. By relying on the external display, errors could be corrected without much memorization. In contrast, subjects in the high-cost condition relied more on their memory to keep track of their task as well as to detect and correct errors. It was also shown that over practice, subjects in the high-cost condition did manage to reduce the number of uncorrected errors.

Conclusions and Discussions

The results support two predictions derived from Anderson's (1990) rational analysis framework. Given an assumed constant expected gain, when the cost of accessing information in-the-world increased, the cost of a perceptual-motor strategy becomes greater than the cost of a memory strategy. Under such conditions, the *optimal* strategy is to encode task-relevant information from the external environment into working memory, thereby reducing reliance on the external environment. The results indicate that when the cost of accessing external information is high, people invest more time storing information in their internal environment and rely less on the external environment to do the task. In contrast, when the cost of accessing external information is low, people spent less time encoding and rely more heavily on the external environment.

Our second prediction was upheld as well. In the high-cost condition this switch to the memory strategy was made despite its higher error rate. Indeed, the decrease in the number of uncorrected errors indicates that with practice our subjects became better at detecting and correcting errors. This finding suggests subjects were *optimizing* the strategy to reduce the overall effort required to do the task.

Under the rational analysis framework, cognition tends to optimize performance by balancing costs and benefits in a given information processing task. Our results show that the cost of information access could induce a switch from reliance on information in-the-world (perceptual-motor strategy) to in-the-head (the memory strategy). We found that although memory is a limited resource, there are conditions under which people can use it to *optimize* performance.

Acknowledgments

The work reported was supported by a grant from the Air Force Office of Scientific Research AFOSR#F49620-97-1-0353 as well as by the National Science Foundation (IRI-9618833).

References

Altmann, E. M., & Gray, W. D. (1999). Preparing to forget: Memory and functional decay in serial attention. *Manuscript submitted for publication*.

Altmann, E. M., & Gray, W. D. (2000). Managing attention by preparing to forget. Proceedings of the *Human Factors and Ergonomics Society 44th Annual Meeting* (pp.). Santa Monica, CA: Human Factors and Ergonomics Society.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Lebiére, C. (Eds.). (1998). *Atomic components of thought*. Hillsdale, NJ: Erlbaum.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, *7*(1), 66-80.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, *46*(3), 461-488.