LOST IN SPACE: THE SURPRISING ROLE OF INFORMATION SPATIAL LAYOUT

by

Jooyoung Jang

B.A. in English Language & Literature, Psychology, Yonsei University, Seoul, 2005

M.A. in Psychology, Yonsei University, Seoul, 2007

Submitted to the Graduate Faculty of

Arts and Sciences in partial fulfillment

Of the requirements for the degree of

Master of Science

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This thesis was presented

by

Jooyoung Jang

It was defended on

October 31st, 2008

and approved by

Christian D. Schunn, PhD, Associate Professor

Stephen C. Hirtle, PhD, Professor

Timothy Nokes, PhD, Assistant Professor

Thesis Advisor: Christian D. Schunn, PhD, Associate Professor

LOST IN SPACE: THE SURPRISING ROLE OF INFORMATION SPATIAL LAYOUT

Jooyoung Jang, M.S.

University of Pittsburgh, 2009

Prior research has found that information presentation formats matter for how easily people understand certain information and solve problems using the presented information. The major finding from those studies is that information should be presented in a type of format that cognitively best fits to characteristics of given problems. Much of this prior has focused on conceptual elements of presentation format (e.g., words vs. diagrams, or graphs vs. tables) rather than more physical elements. However, with effort and strategy considerations in mind, more physical elements may also influence performance. Here, I focused on an understudied physical element that is pragmatically important and potentially theoretically exciting: the spatial layout of the information. Specifically, is there a difference between superimposed information (e.g., presented in a pile of pages) and distributed information (e.g., when the same pages of information is spread out sticking on a wall)? This question originated from an observation of meteorologists making weather forecasts. In an earlier study, meteorologists made a forecast in two conditions: mapwall and computer. Although the computer user could use animations and comparisons, there was no difference in accuracy but there was a large, nearly 50% time difference: Mapwall users made predictions far faster than computer users. The purpose of this master's project was to develop a lab task to replicate the effect and to reveal the underlying mechanism. The results showed a large speedup effect of the distributed format, but only on a task involving information integration. There, distributed task speed was almost twice as fast. An underlying mechanism for the effect, the strategy selection hypothesis, was also tested.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

The way information is presented has been an issue in many domains such as science learning ([14]Larkin & Simon, 1987; [18]Shah & Hoeffner, 2002) and decision-making ([5]Card et al., 1999; [6]Chen, 1999; [20]Tufte, 2001; [23]Ware, 2000). To facilitate learning and problem solving, those studies have used tables, graphs, and diagrams, showing different effectiveness in various problem situations. Although this work generated as many particular suggestions as the number of studies done, there is a converging thought that performance improves when the information presentation format cognitively fits with the mental representation a problem required ([21]Vessey, 1991). However, previous studies were done with information structure / format per se (e.g., whether they are in tables or graphs), and little research was focused on more purely physical factors like the spatial layout of information.

For physical factors, fit to mental representation is not likely to matter. Thus, it might seem that spatial layout would make no difference in performance when the given information is same. But there is a study showing a fairly good possibility that it is important and does have an effect. Trafton and his colleagues (2007, unpublished research) investigated complex visualization usage by intermediate meteorological forecasters. They observed an anomaly that there was a difference concerning the time people spent to complete the task on the computer and the mapwall. Even though people in mapwall condition could not use animations and map comparison tools commonly used by forcasters to improve forecasting, they made predictions far faster than people in the computer condition, with no difference in accuracy. Specifically, the time spent in the computer condition was almost double that in the mapwall condition. As the difference between the two conditions was primarily in the spatial layout of the same information, it is necessary to delve into the role of spatial layout of information in problem solving. It will be useful to also show that the paper vs. computer difference in conditions of that experiment is not what produced the results.

1.1 Spatial layout of information and the nature of task

There are times that we can remember where the information is but not what the content is about or sometimes both. This memory has been called *incidental memory* for location of information. Rothkopf (1971:[17]) empirically demonstrated that people formed this kind of memory while they were reading a text. He gave participants 12 pages of passage to study. Immediately after they finished reading, participants were asked to answer questions about the content and to recall the location of the information on which each question was based. The results showed that readers recalled locations of information better than chance and accurate recall of the contents and the locations were correlated. This study suggests that spatial location of information is encoded in association with the content, which in turn suggests the potential value of spatial layout of information. It also seems that the process is somewhat automatic since participants were not asked to memorize locations initially.

While Rothkopf investigated the simple relationship between memory for location and the content, there is a recent study supporting the claim that spatial layout matters when people perform complex integration (Vincow & Wickens, 2007:[22]). In the study, participants were given tables of information and asked to perform simple or complex integration (division, multiplication) between pieces of information in the tables. Sometimes information required for integration was located within a table panel, but sometimes information was separated between panels. It was found that people performed poorly when they had to integrate pieces of information separately located between panels, but this effect occurred only when they completed a complex integration task. This interaction result suggests that the influence of spatial layout can be differentiated by the types of task, which may relate to the general idea that the information presentation format should cognitively fit with the mental representation a problem requires. In the aspect of underlying mechanism, this finding implicates different access rate to information in working memory caused by spatial separation of information.

One possibility to explain these findings can be found in a concentric model of working memory proposed by Oberauer (2002:[15]). He conceptualized working memory as *a concentric structure of representations with three functionally distinct regions: the activated part of long-term memory, the region of direct access, and a focus of attention* (Figure 1). In his study, participants were given two rows of digits which should be memorized and potentially updated by upcoming arithmetic operations, but not only one row requiring updates. Either one of the two rows were displayed in red which indicated that the digits in that row are "active" and these digits should be updated, while the digits in the other row are "passive". The experiment results suggested that the digits in the active row were held in the region of direct access, whereas the digits in the passive row were held in the activated part of long-term memory.

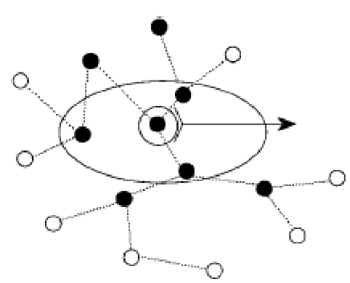


Figure 1. Adapted from Oberauer, 2002. A concentric model of working memory. Nodes and lines represent a network of long-term memory representations, some of which are activated (black nodes). A subset of these items is held in the region of direct access (large oval). Within the region of direct access, one item is selected for processing by the focus of attention (small oval). Activated items outside the region of direct access

form the activated part of long-term memory, which is accessible only indirectly via associative links (dotted lines) to representations in the more central regions

Based on this account, it can be hypothesized that spatial layout can play a role of determining which information is held in the region of direct access. When pages of information are spread out and placed in fixed locations, their location may work as a spatial cue declaring them as active which may make the information stay in the region of direct access. Whereas, when the information is buried in a pile of pages, no such spatial cue exists and therefore the information is placed outside the region of direct access.

1.2 Strategy

The possibility of spatial cuing can be regarded as a bottom-up explanation because the access rate to information is determined by the availability of spatial cues in the problem, not by the problem solver. However, it is important to consider a possibility of top-down effect since people can adopt variety of different strategies while doing problem solving. Recently, researches have suggested that people choose an optimal strategy to reduce the burden of internal memory ([3]Ballard, Hayhoe, & Pelz, 1995; [11]Fu & Gray, 2000; [12]Gray, Schoelles, & Sims, 2004). When information is *in-the-world* and the cost of accessing external information is low, people do not bother to memorize information and employ a perceptual-motor strategy. They return to necessary information with a saccadic movement or head turn. In contrast, as the cost to access increases, information *in-the-head* becomes relatively more accessible and people tend to employ a memory strategy. That is, people spend more upfront time to store or encode information into their internal memory, thereby avoiding high-cost external information.

This perspective can offer an alternative explanation about the effects of spatial layout of information. When the pages of information are spread out, presumably it constructs a low-cost external environment because information is a head-turn away. Therefore, people choose the perceptual-motor strategy and rely on in-the-world information. Whereas, when information is presented in a pile of pages, the cost to access external information increases because people must turn over the pages to find information in need. This differential access cost might make the memory strategy the more optimal solution.

To address these issues, I presented two studies. The first study replicates the spatial layout effect in the lab. The second study provided a test for why the effect occurs.

2.0 STUDY 1: DEVELOPING A LAB TASK

Developing a lab task was necessary for two reasons. Practically, making a weather forecast requires highly specialized training and experience, which makes natural participants rare and expensive. Although it has its value as a real-world task, it is hard to implement in the lab. It is time consuming both for researchers and participants to learn even the basic principles about weather forecasting and apply those to generate a task or solve the task. More importantly, there is no need to adhere to the weather forecasting task because of the following reason. Theoretically, to give external validity to the particular observation, there is a need to show that it is a general phenomenon. If it is not a one-time special case and has a certain consistent and meaningful underlying mechanism, the effect should be replicable in another task.

Although hopefully a general phenomenon, it is likely that a new task must have some essential features residing in the weather forecasting task to replicate the layout effect. The problem space in weather forecasting task consists of many pages of information, presenting meteorological elements such as temperature, air pressure, the direction of the wind, and the velocity of the wind. Each information type is also presented for many time points, since forecasting is a job to make inferences by analyzing meteorological elements unfolding over time. This problem space yields the most important feature of the task: it requires information integration across dimensions and over time. Since meteorologists give us relatively simple predictions like temperature and chance of rain, they must have interpreted and integrated information in some meaningful way based on the prediction rules they have learned. On the basis of this speculation, a simplified task that requires information integration was developed and tested.

2.1 Deer task

The deer task asks participants to compute out how many adult deer and fawn survive after a period of time in each of five regions by applying prediction rules. Participants were told that deer die when they lose too much weight and, extreme temperatures and lack of food cause the weight loss. Prediction rules consist of temperature, food (snow depth), and total weight loss rules. The rules are presented in text with a diagram (APPENDIX A. Summary; A diagram of rules for prediction). The season for making predictions is always winter, and it is supposed that adult deer and fawns lose different amount of weight under each of the weather conditions. The rules are stated as follows:

> **TEMPERATURE**
>
> In winter, since it is cold, deer usually are confronted with "**Cold Stress**." Even though they grow a thick winter coat, they start losing weight under extreme temperatures (- 4F or less). Weight loss for adults and fawns are 5lbs and 7lbs respectively during a 2-week period if the temperature is that low. Ironically, deer can also suffer from "**Heat Stress**," thanks to their winter coat. When there are warm winter days ranging between 30F and 45F during a 2-week period, adult deer lose 3lbs and fawns lose 5lbs.
>
> **FOOD**
>
> The supply of winter food is closely related to "**Snow Depth**." When it snows heavily and the snow depth is 7 inches or more, deer tend to yard up in a smaller area, which can help them to avoid cold stress. However, it not only narrows down the range of food searching area but also causes more energy loss during their quest. Consequently, snow depth over 7-inches yields weight loss of 7lbs for adults and 10lbs for fawns in a 2-week period.

TOTAL WEIGHT LOSS and its effects on **SURVIVAL**

Total weight loss of less than 30lbs has no survival impact. Over 30lbs weight loss leads to only 80 out of 100 adults and 60 out of 100 fawns to survive in any region. After that, the survival chance for both groups cut down by 50% for every 10 lbs weight loss from the previous weight loss, for example, with the first 35 lbs loss, the 100 adults will drop to 80 and then10lbs more loss (45 lbs total), only 40 adults will survive.

After these rule pages, there are pages of slides containing information of temperature and snow depth for ten different days (Two pages for each day, 20 pages in total). The information of temperature and snow depth is presented in a paired manner according to a date (Figure 2). Because there are five regions, and predictions are to be made separately for adults and fawns in each region, participants need to make ten pairs of predictions for weight loss on each of the days, and then for resulting deer populations.

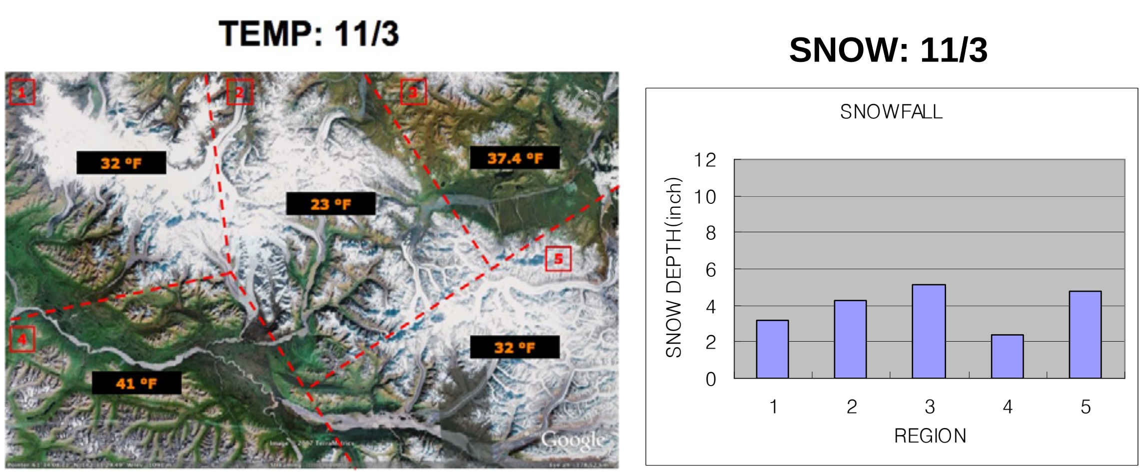


Figure 2. Two sample pages of information of temperature and snow depth

Specifically, on each day, participants must calculate how much weight loss is caused with adults and fawns, respectively and add up those numbers as days go by while applying total weight loss rules when it gets over 30 pounds. The initial population is 100 adults and 100 fawns.

2.2 Method

2.2.1 Participants

Twenty-nine undergraduates from the University of Pittsburgh participated in this study for course credit and were randomly assigned to one of the two experimental conditions.

2.2.2 Design and procedure

The two conditions were designed to differentiate spatial layout of information. In the computer condition, pages of information could only be seen one at a time. Participants could browse through different pages of information by clicking table of contents placed in the left frame of the window. In the wall condition, pages of information were spread out and stuck on a wall (Figure 3). There were 16 participants in the computer condition and 13 in the wall condition.

Participants in both conditions were first led through a practice session by an experimenter. The practice session was done on the computer. Participants were first instructed to read the rules carefully and told that they could refer to the rules whenever they needed them. Then, they went through example pages of temperature and snow depth with an experimenter and tried to apply the rules and compute out how much weight loss occurred in each region. The practice session continued until they could use the rules correctly.

To assist calculation and make participant solution methods visible, a clipboard and paper were provided. The clipboard pages had lines distinguishing five regions as drawn in a temperature slide. No restrictions were given on how to use the notes.

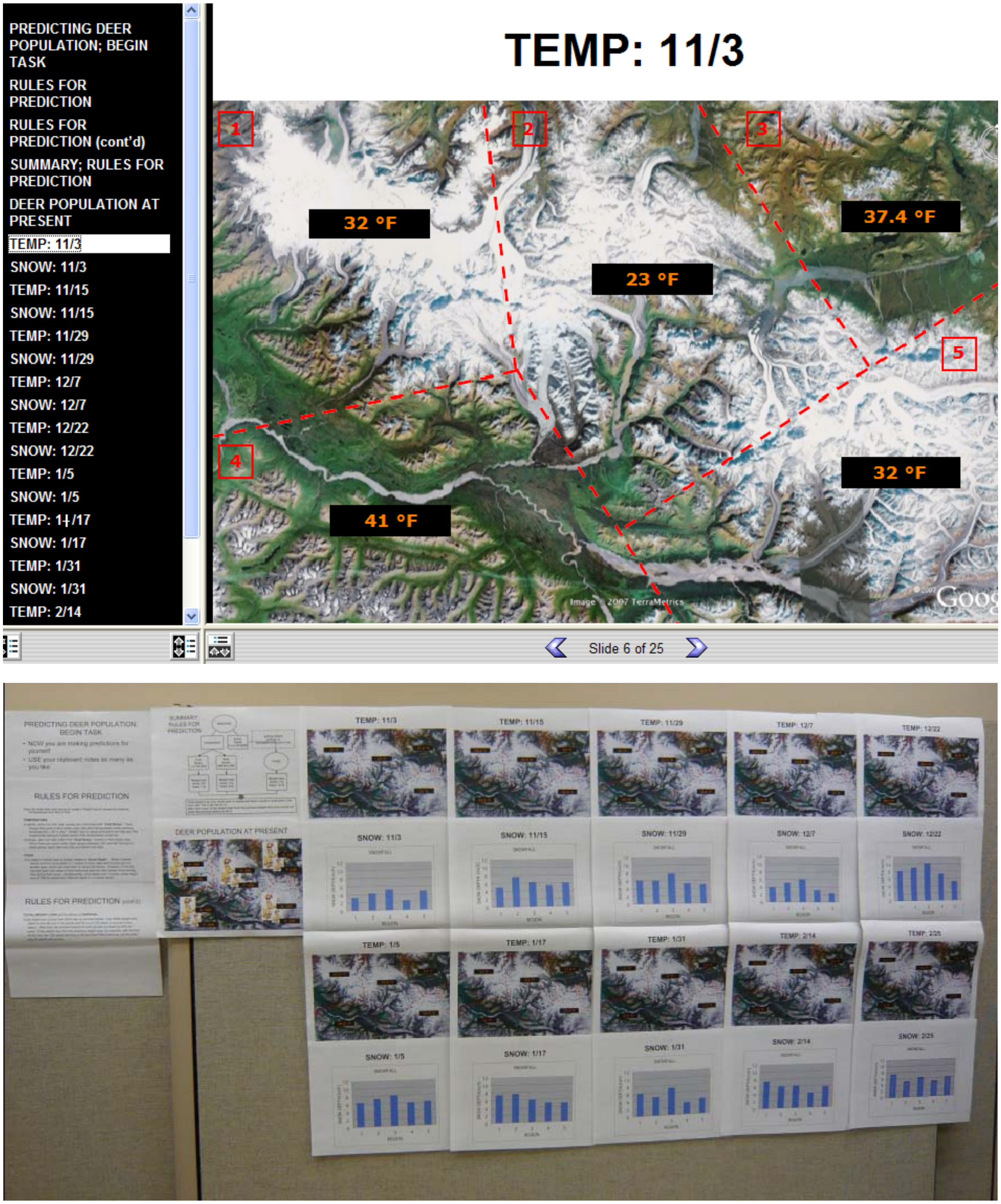


Figure 3. Spatial layout of the computer condition (upper panel) and the wall condition (lower panel)

2.3 Results and Discussion

Three participants who did not finish the experiment in the given time (1 hour) were excluded from analyses. Outliers were defined as accuracy lower than 6 out of 10, which was one standard deviation below the mean. Five participants were excluded as outliers. Analysis of variance (ANOVA) of time for condition showed a significant effect of condition ($F(1, 20) = 5.8$ $p < .05$), with participants in the wall condition performing significantly faster than those in the computer condition. As shown in Figure 4, participants in the wall condition spent less 6.9 minutes less time, that is, a 32% time difference. There was no significant difference in accuracy or practice time. Thus, the deer task nicely replicated the prior spatial layout results, establishing the generalizability of the effect across tasks and subject populations. Therefore, study 2 can now focus on establishing possible causes of the effect. Note that study 1 still confounded spatial layout with information medium (i.e. paper vs. computer). Study 2 will remove this confound.

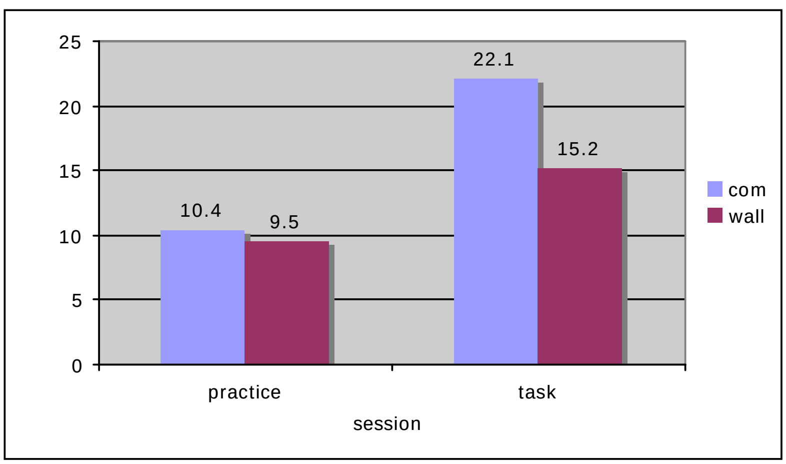


Figure 4. Mean times for computer and wall condition in practice and task session

3.0 STUDY 2: SPATIAL CUEING VS. STRATEGY

3.1 Method

3.1.1 Participants

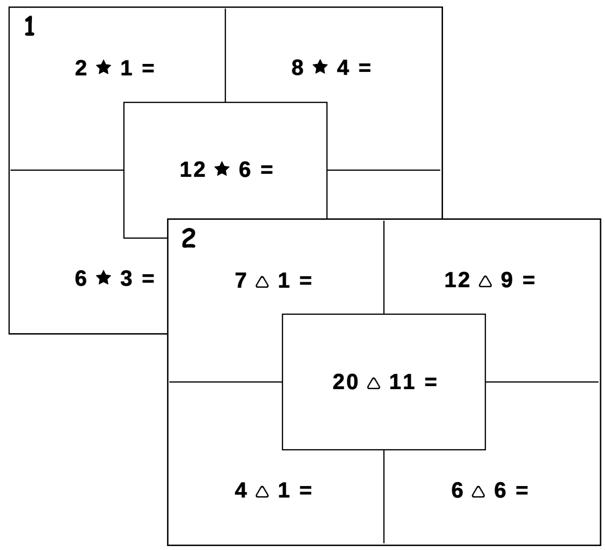
Eighty-three undergraduates (41 women and 42 men) at the University of Pittsburgh participated for course credit. Ten of these(5 women and 5 men) participated in the video taping session, they were native English speakers. Participants were randomly assigned to one of the four experimental conditions, but video was only gathered from the integration task conditions.

3.1.2 Materials

3.1.2.1 Integration task Deer task was used for integration task, which was a same version in Study 1.

3.1.2.2 Non-integration task To examine the generalizability of the layout effect and to rule out the possible effect of having to stand vs. sit on task completion time, a task that does not require any integration process was added. For the non-integration task, symbol algebra problems were developed based on Blessing and Anderson (1996: [4]). To keep the overall structure similar to the deer task, rules were presented on the first page, and 20 pages of arithmetic problems followed. There were five rules, roughly matched to rules in the deer task. The first four rules defined each of the four symbols' meaning in arithmetic functions; A ☐ B = A x B + 5, A △ B = (A – B) / 3, A ⦿ B = (A + B) x 7, A ★ B = A / B x 9. The last rule described the decision criterion which allowed people to judge which problems they should solve and which ones they should not: “Only solve problems if one number is greater than or equal to 10”.

Table 1. Rules and Example Pages for the Non-integration Task

Rules	Example Pages
■ A □ B = A x B + 5	
■ A △ B = (A – B) / 3	
■ A ⦿ B = (A + B) x 7	
■ A ★ B = A / B x 9	
■ Only solve problems if one number is greater than or equal to 10. If both numbers are less than 10, skip that problem.	

Sample pages of the task are shown in Table 1. To match the task load between the integration and non-integration tasks (i.e., deer and symbol tasks), each page was composed of five problems having only one symbol operator, so that only one rule could be used per page. The number of problems to be solved was the same as the number of weight losses to be detected in the deer task.

3.1.2.3 Memory test For testing between the different predictions of the spatial cueing and strategy hypotheses, a memory test was conducted. Participants were asked to recall all the information they saw in the last four pages of the given task. In the integration task, they were to recall the temperature in each region for the last 2 days and draw a graph of snow depth with the inch of snow depth for the last 2 days. In the non-integration task, participants were also asked to recall all the symbol problems they saw, whether solved or not, for the last four pages.

3.1.2.4 Survey A short survey was given to examine the strategy selection hypothesis. There were 12 items asking subjects to think about perceived information availability and their use of an in-the-head memory strategy (i.e. whether they tried to memorize information). For responses, the typical five-level Likert scale ranging from ‘Disagree’ to ‘Agree’ was used. The full version of survey questions are presented in APPENDIX B.

3.1.3 Design and procedure

A 2x2 between subjects factorial design was used with the types of spatial layout (paging or wall) and the types of task (integration or non-integration). In the paging condition, pages of information were put in a binder and each page was attached to cardboard so that it could not be bent, preventing participants from seeing information other pages. Indexes on each page were provided as it was done with the table of contents in the computer condition. Since theoretically the time effect observed between the computer and the wall condition originated from different spatial layouts, not from the medium (i.e. computer or paper) itself, the computer condition should not be different from the paging condition in terms of cognitive processing. The paging condition was adopted to clarify this issue, thereby establishing spatial layout of information as the substantial topic.

Participants were randomly assigned to one of the four conditions (paging-integration, paging-non-integration, wall-integration and wall-non-integration). First, all participants filled out a form regarding their major, academic year and SAT scores. During the practice session, they learned the rules and solved sample problems with an experimenter until they could understand and apply the rules correctly, and then they moved on to the task session. All participants practiced with the computer as a natural training method across conditions or possibly biased towards the paging condition. As soon as they finished the task, the memory test was given without prior notice. Paging people remained seated throughout the whole session, but wall people were instructed to stand up during the task session.

3.2 Results and Discussion

3.2.1 Time and accuracy: effects of spatial layout in association with the types of task

The time spent to finish the task was submitted to analysis of variance (ANOVA) with two factors: type of spatial layout (paging or wall) and type of task (integration or non-integration). The means and standard deviations for practice time, task time, and accuracy of integration and non-integration task are presented in Table 2. Participants who failed to finish the practice or the task within two standard deviations above the mean time were not included in the analyses. This resulted in a loss of 5 participants; 4 from the integration and 1 from the non-integration tasks. In total, data collected from 78 participants were subjected to analyses; 39 in integration task with 19 men and 20 women, 39 in non-integration task with 21 men and 18 women.

The time effect between paging and wall condition was observed as expected, which confirmed that the paging condition is not different from the computer condition. Therefore, the medium delivering information is not a critical factor. However, as it is presented in Figure 5, the time effect was observed only in the integration task, F (1, 38) =32.7, p < .01, Cohen's d = 1.8.

Participants in the paging condition spent about 6 more minutes to finish the task than did the wall subjects. In the non-integration task, both wall and paging people spent similar amounts of time to solve the symbol arithmetic problems.

The accuracies of both conditions showed no difference within either type of the task. There was a marginally significant difference on practice time between the two types of spatial layout in the integration task, (F (1, 38) = 4.1, p = .05). However, as it is shown in Table 2, paging people took longer to finish the practice, and even if people in the paging condition spent a little more time on practice, they still spent more time to solve the task. To ensure task time effect with the influence of practice time, an analysis of covariance (ANCOVA) was done. Task time was significant even if practice time effect was included (F (1, 38) = 16.2, p < .001).

Table 2. Means and Standard Deviations for Practice Time, Task Time, and Task Accuracy of Integration and Non-integration Tasks

	Integration (Deer Task)				Non-integration (Symbol Task)			
	Wall		Paging		Wall		Paging	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Practice time (min)	11.8	1.8	13.1	2.0	4.6	0.6	4.9	0.8
Task time (min)	10.6	2.7	16.8	4	8.3	1.6	8.3	1.5
Task accuracy (%)	75.5	16.5	79	13.3	95.7	3.5	94.9	5.0

In the integration task, there was no other significant difference observed at practice time, task time or task accuracy by gender, academic year, major, SAT (math) score, or their interactions. In contrast, for the non-integration task, there were one significant and a few marginally significant differences between groups by gender and major. Male participants showed significantly higher task accuracy as compared to female participants (F (1, 38) = 7.0, $p < .05$), and they tended to finish the task about a minute faster than females (F (1, 38) = 4.1, $p = .05$). Science related majors solved symbol arithmetic problems marginally faster and more accurately than other majors (F (1, 38) = 3.1, $p = .09$; F (1, 38) = 2.9, $p = .10$). These tendencies are likely caused by the mathematical nature of the symbol task.

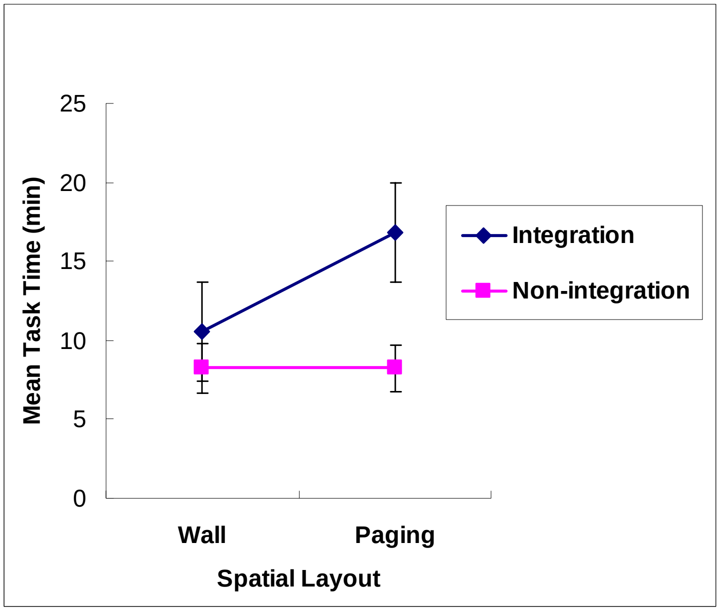


Figure 5. Mean time and SE bars to solve the task for each task in each spatial layout condition

3.2.2 Memory test: spatial cueing vs. strategy

The score of correctly recalled information was computed according to the following coding scheme. Having a wrong temperature or wrong snow depth in the correct place was meaningless. Each correct value written in any place within the page was counted as one point. A correct value in the correct location counted as two points, which would result 40 points in total. However, there was no such case and that simplified the coding scheme to be giving one point per correct value within the given page. The maximum score per participant was then 20, five points for each of the last four pages. In the non-integration task, although there were three units of information to remember (i.e. left operand, operator, and right operand), a problem was counted as a whole as one point. Since the same symbol was used within a page, recalling the correct symbol on a page has little meaning.

The spatial cueing hypothesis explains the time effect by availability of spatial cues in physical space, which holds information in the region of direct access. If this hypothesis is supported, the memory test should show higher scores in the wall condition, regardless of the type of task. Only the wall condition has spatial cues provided from fixed location in the wall space, and those cues make information active within the boundary of direct access. Therefore, it was expected that people in the wall condition would recall information accurately. This result should be yielded equally whether they did the integration task or not because presumably the process of spatial cueing occurs in an automatic and bottom-up manner.

If the strategy selection hypothesis that people use information on the wall as in-the-world memory is supported, higher accuracy in the memory test should be shown only in the paging-integration condition. Since in the paging condition the cost to access information is high, people are expected to deploy in-the-head storage to reduce the frequency of turning over pages. Although the paging condition requires high cost to access information in both types of task, higher accuracy in the non-integration task is not expected because people would not need to look up previous page information when they are not required to integrate information. The expected schematic graphs from the spatial cueing and strategy selection hypotheses and the graph representing the actual results are shown in Figure 6.

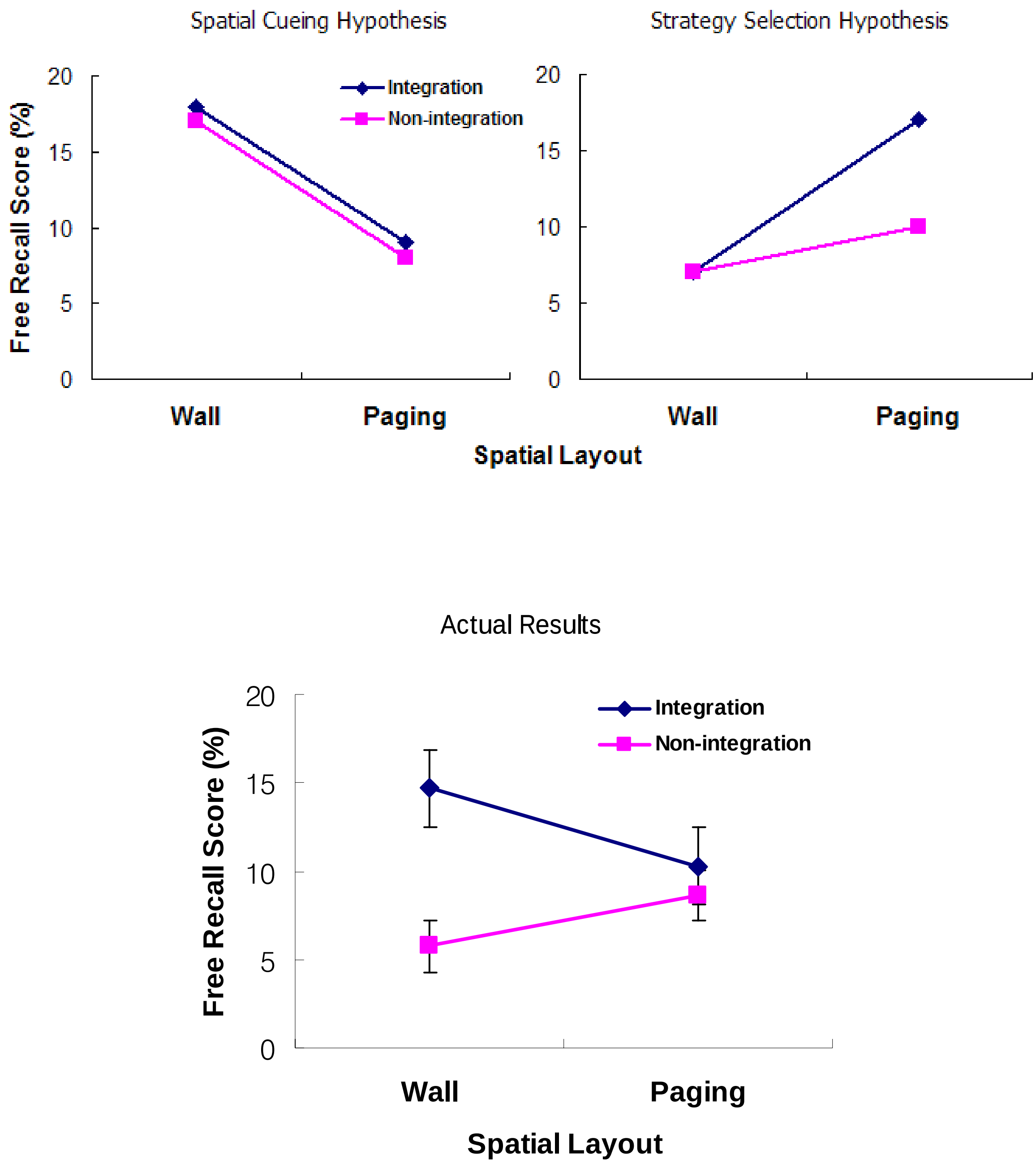


Figure 6. Mean Recall Scores by Condition for the Expected Results from the Two Hypotheses and the Actual Results with SE bars

The actual results overall support neither hypothesis. The wall subjects recalled significantly more information units than the paging subjects, F (1, 33) = 8.15, $p < .01$, as the spatial cueing hypothesis predicts. However, this trend was observed only with the integration task unlike the assumption of the spatial cueing hypothesis that predicts high free recall scores of wall subjects for both types of task. Although the non-integration task data shows the opposite trend, which agrees with expected results from the strategy selection hypothesis, the trend was not statistically significant.

To further examine memory performance, a forgetting function graph was drawn according to page numbers. This revealed different forgetting patterns according to the type of task (Figure 7). The non-integration task, which requires simple arithmetic calculated separately on each page, showed a recency effect generally found in immediate recall tests. In fact, very little was remembered from any page except the last page in the non-integration task. However, in the integration task, people recalled far less information from the last page than they did from the fourth to the last page. They also showed significantly higher memory for temperature pages ($F\ (1, 155) = 39.7$, $p < .001$), which partly can be explained by the fact that temperatures were easier to remember since they were composed only of integers, whereas snow depths were expressed with one decimal and people might direct less attention to the numbers after the decimal point in this task. In addition, all snow depth information was presented with a graph, and the rule of importance was whether the bar of snow is over a certain point (i.e. 7-inch) or not. Therefore, people might not have to encode the exact snow depth value to detect its effect on weight loss. Even considering these complications in the stimuli, the forgetting function clearly confirms that a different kind of cognitive processing is operating during the integration task, since subjects recalled significantly more items from the fourth-to-last page than from the second-to-last page ($F(1, 77) = 5.2$, $p < .05$). Those two temperature pages had the same number of units to be remembered (five) and the same number of one-digit temperatures and two-digit temperatures (three and two, respectively). Therefore, this surprising forgetting function is evidence that the nature of integration tasks is fairly different from non-integration tasks vis-à-vis incidental memory.

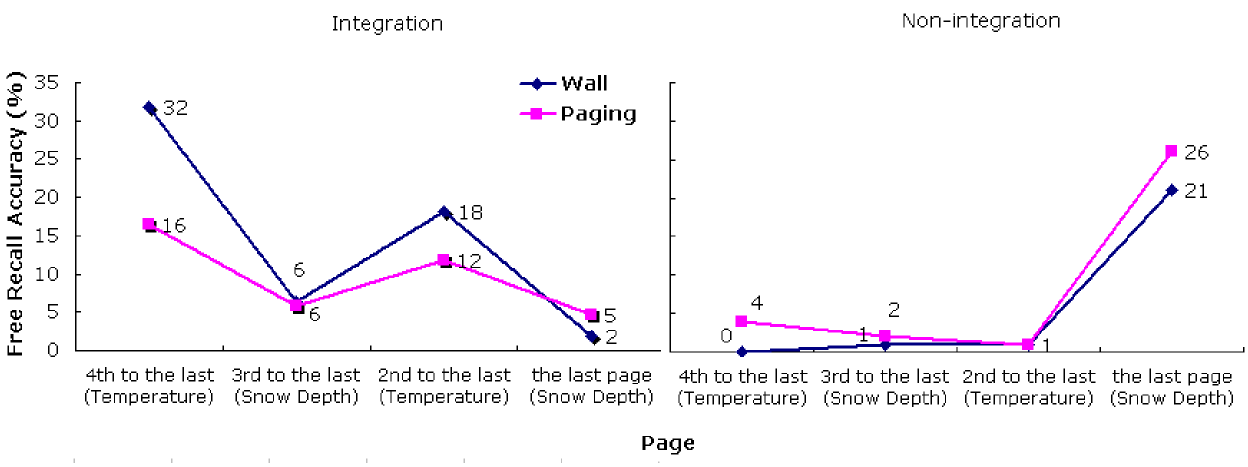


Figure 7. % recall accuracy as a function of recency in both condition for each task

3.2.3 Survey responses

The survey was conducted to investigate additional differences between the wall and paging conditions at the conscious level. It did not yield many interesting results (see Table 3). There were several significant differences by task type: participants who worked on the non-integration task reported more trouble finding and remembering information from prior pages (Questions 2, 3, 5, 6, and 9 *p*s $< .05$). These differences might reflect the nature of the non-integration task, since it was composed of many repetitive symbols and numbers having no relationship among them, which naturally causes a lack of coherence. However, it is noteworthy that people who solved the integration task did try to keep temperatures and snow amounts in their head. This confirms that people underwent different types of cognitive processes in the different types of task. As a whole, the surveys shows that the time and cognitive processing difference by formats seems subtle enough to be reside below consciousness, which suggests focusing more on implicit methods, possibly eye-tracking for a next study.

Table 3. Mean Ratings for Selected Survey Questions. The left-most column indicates the actual order of items used in data collection. The survey was 1 to 5 likert scale labeled as disagree (1) to agree (5)

		Integration Task		Non-integration Task	
		Wall	Paging	Wall	Paging
1	Before the task started, I thought to write down the rules in my notes	2.4	2.5	2.9	2.2
3	I generally relied on information being available on paper instead of trying to remember it	3.7	3.5	4.6	4.8
7	I purposely tried to keep temperatures (numbers) and snow amounts (symbols) in my head	2.5	2.3	1.4	1.3
5	All the information from different pages blurred together	2.7	2.7	3.3	3.5
8	I got information from one page confused with information on another page	2.6	2.1	2.9	2.9
9	I got lost about where I was in the task (e.g. what page)	1.9	1.6	2.5	2.8
	I had trouble				
2	Remembering information from prior page and often had to look back	3.4	3.3	4.5	4.5
4	Finding information I need	1.6	1.1	1.5	1.2
6	Remembering the rules	2.2	2.3	3.5	4

3.2.4 Time log data

The time log data are logs of participants' problem solving progress: which information page participants are looking at what time, for how much time, and whether they go back and forth between pages. I recorded time logs of five participants in each of the four conditions by hand. Additionally, ten more subjects of time log data were collected from video data recorded in the integration task; five in the wall and five in the paging condition. Thus, in total, there were 10 time logs collected in the non-integration task and 20 time logs in the integration task. However, 4 participants in the integration task who spent more time than two standard deviations above mean task time were excluded.

From this data, I counted how many times subject went back to the rule pages. In the non-integration task, there was a marginally significant trend that paging subjects (M= 15.2) referred to the rule page more than wall subjects (M=11.4), F (1, 9) = 3.9, p = .08. In the integration task, paging subjects went back to the rule page significantly more frequently than wall subjects, F (1, 15) = 6.2, p < .01. The mean rule referring frequency of paging subjects (13.3) was about four times that of wall subjects (3.83). As it is shown in Figure 8, paging subjects consistently refer to the rule pages more frequently through out the task.

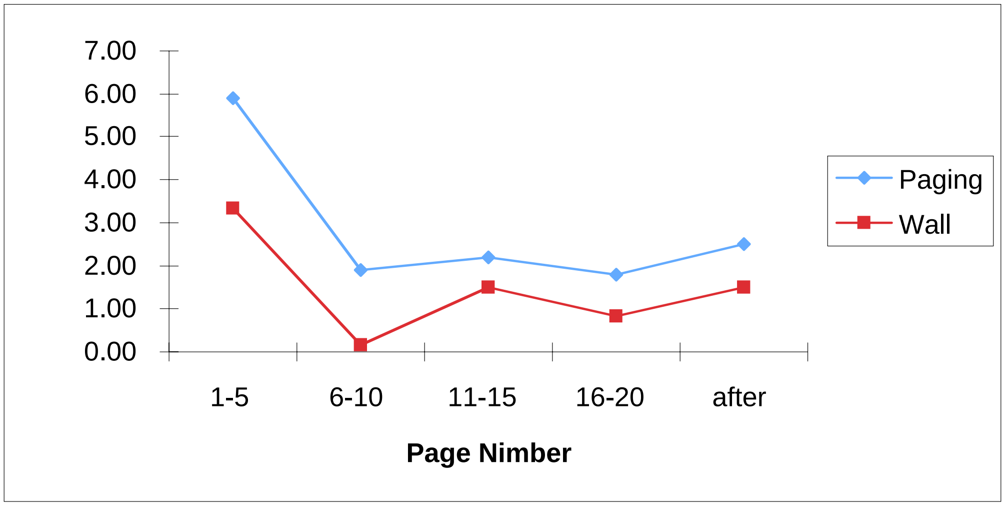


Figure 8. Mean Rule Frequency Split by Page in the Integration Task

Both wall and paging subjects went back to the rule pages most during the first five pages, which is expected, since participants were not familiar with the rules at first. From page 11 to 15, there was a slight rise in the frequency, which may have resulted from the design of the deer task; the first half of the 20 pages were about heat stress and the remaining half were about cold stress. So starting from page 11, participants needed to go back to the rule page and re-familiarize themselves with the cold stress rule. After they were done with the problem pages, participants had to refer to the rule pages once more, for making a final prediction using the total weight loss rule. This provides an explanation for another slight rise in the frequency during 'after' pages.

The time log data also allows an analysis of participants' pace in turning pages, which can help to further localize the time effect. In other words, the question is where the paging people spent extra time on and in what pattern. To answer this question, a graph of processing pace was drawn based on median times of reaching to each page (Figure 9). The median times were used to prevent the graph from being affected too much by few saliently fast wall subjects and slow paging subjects, because the N for each condition is 10, which is rather small. In the non-integration task, there was no difference in terms of problem solving pace between wall and paging subjects. However, in the integration task, there was a clear trend that paging subjects started off at a similar pace as the wall subjects for the first three pages but then gradually separated to finish with a large time gap between the two conditions.

To explore further the nature of the difference between wall and paging subjects, data of 10 subjects was selected with task time within one standard deviation from the mean, among the 20 subjects providing the 10 hand-written time logs and 10 videos. The task time ranged from 8 to 12 minutes for wall subjects and from 12 to 20 minutes for paging subjects. This data selection was expected to maximize differences between conditions and produce clear process results.

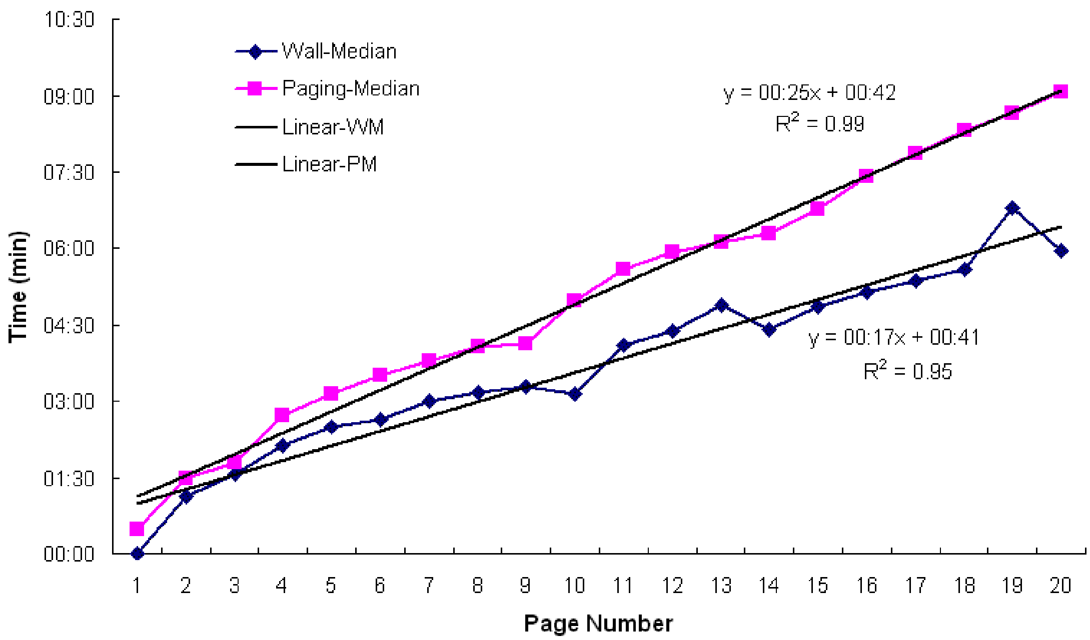


Figure 9. Patterns of Processing Pace in Integration Task. Points represent medians time that participants reached each page for the first time

Five subjects were selected from each condition, and their task times were broken down by two phases. One is the time to reach the last page, which indicates the time spent to process all the information pages and come to the last page. The other is the final integration time showing time spent after finishing with the last page until task completion. As the graph shows in Figure 10, three out of five wall subjects finished the last page earlier than all paging subjects. These three subjects are consistent with the overall trend observed in the time log data. In the final integration phase, all the paging subjects spent more time on computing up the numbers and making predictions than the wall subjects. Even though two wall subjects took a similar amount of time to reach the last page as the paging subjects, both were far faster at the final integration phase than other people, suggesting that they had also had a processing benefit earlier on, but had chosen to complete more of the task along the way.

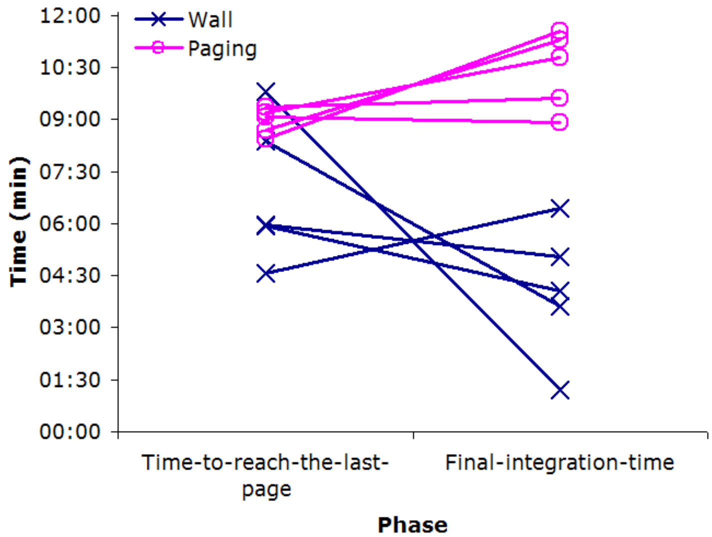


Figure 10. Time log trend for each participants broken down into the time to reach the last page and the final integration time

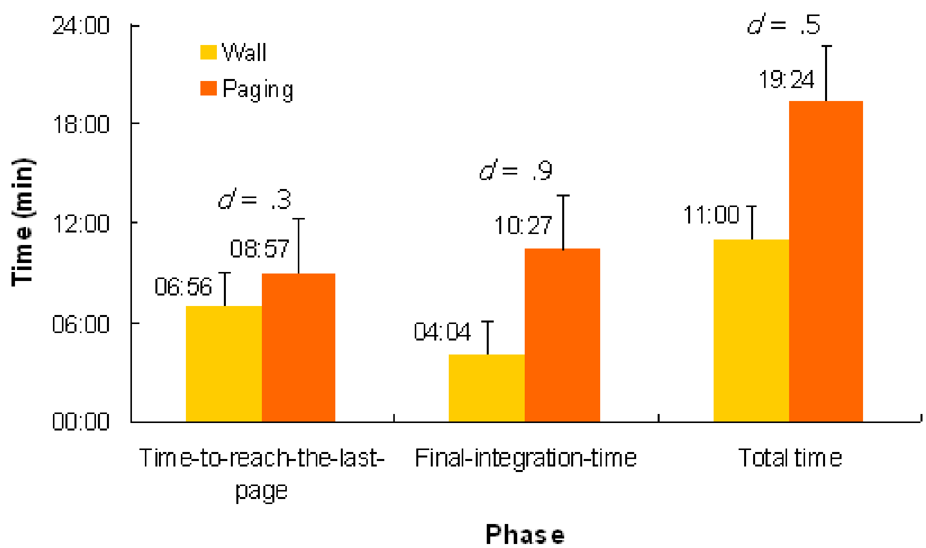


Figure 11. Mean Task Time Split by Phase. The first two bars show mean time spent until people reached to the last page for the first time. The middle bars indicate mean time took for after calculation, that is, the time between when people reached to the last page for the first time and when they finish the entire task. The right bars represent mean time spent to finish the entire task (i.e. sums of the left bars by the condition, respectively)

The overall pattern of time consuming of the selected ten subjects can be summarized with Figure 11. It shows about two minutes difference between conditions until subjects reached to the last page, which might reflect paging subjects' higher frequency of referring to the rules and physical time for turning pages. The most salient difference can be observed during the final integration time phase. There is about six minutes difference, which contributes three fourths of the total time difference (eight minutes). Therefore, this suggests that the spatial layout of information affects not only the process of browsing problem space and gathering information but also the productivity of aftermath information integration.

3.2.5 Verbal Protocols

Among the ten collected video data, only five could be included in the analyses: data of one subject in the paging condition could not be used because she spoke too quietly. Another four data were excluded due to task time, which implied interference of verbalization in the task. Two participants in the wall condition got slowed down, and one of them made a lot of comments explaining the problem solving process and analyzing the problem situation, both of which they were asked not to do. In the opposite direction to the wall subjects, two paging subjects became unusually fast. The task time of those four participants were above or under two standard deviations from mean, so they were excluded and five transcriptions were left to undergo analyses: three in the wall condition and two in the paging condition.

The high rule-seeking frequency of the paging subjects opened the possibility that they might experience uncertainties and feel more uneasiness with losing track of the task relative to the wall subjects. Based on this hypothesis, remarks or actions reflecting uncertainties and coping behaviors were analyzed. The uncertainty remarks or actions included utterances expressing forgetfulness of rules, losing track of what page they were working on, and coping behaviors such as making charts or marking the dates down on the notes. The rule-reading actions were not included, since it would overlap with the rule counts analysis that is already done with time log data. Figure 12 shows how many uncertainty remarks or actions were made by each of the five subjects. While all wall subjects and a paging subject made fairly few uncertainty remarks and actions, the other paging subject showed high level of uncertainty, making some remarks as follows,

> "I'm marking the date down now because I fell like I am going to lose count, or lose track of the date, if I don't do that. So I'm putting boxes around the predictions or the information that I already have. And I'm going to use the rest of these boxes to make dates obvious."

This kind of thinking or planning actions might cost extra cognitive resource and time while they were collecting information from each page until they reach the last one.

The uncertainty remarks and actions partly explain the time difference during the task, which were mostly generated by frequent rule-seeking behavior. In the final-integration-time, paging subjects did look up the rules more frequently. However, this might not be enough to explain the time difference observed during the final-integration-time, since it shows almost 6 minute difference which takes up three quarters of the entire task time difference. There can be at least three hypothetical explanations for this: slower calculation, recalculation, more calculations, or a combination of these factors.

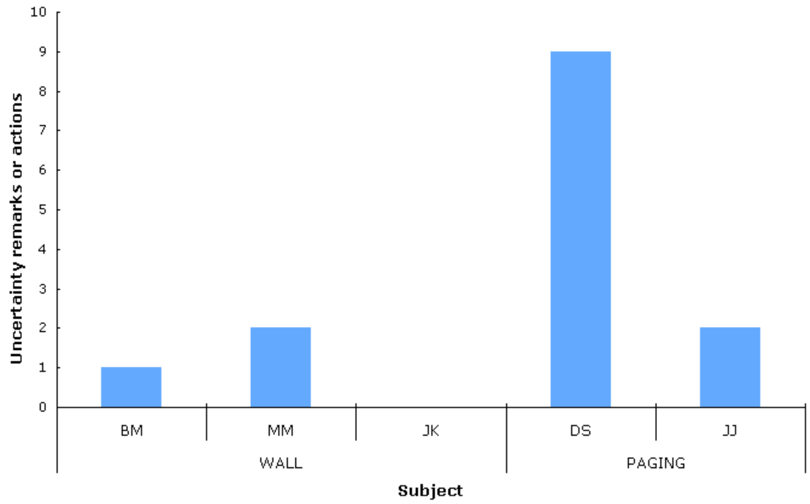


Figure 12. Uncertainty Remarks or Actions

The number of calculations could not be counted exactly because the videos were shot in distance, not close enough to follow their note-taking so that the participants do not feel uncomfortable. Based on the verbal protocols only, there was no salient sign of a small number of calculations among wall subjects. However, this could imply slower calculation for paging subjects under the assumption that both groups had same number of calculations to do. In fact, subject DS in paging condition, who showed the longest final-integration-time among the five subjects, counted and calculated some numbers twice. At present, it is hard to draw a firm conclusion, but uneasiness caused by underlying uncertainty might continue to affect final calculation and yield a time cost.

3.2.6 Correlations between variables

To obtain an integrative picture and clarify relationships between results separately investigated above, correlation analyses on the integration and the non-integration task were conducted. As it is shown in Table 4, several plausible relations were confirmed.

First, task time correlated with its two sub-constructs, time to reach the last page and final integration time, but in different patterns by the nature of the task. In the integration task, task time shows significantly positive relationship with both sub-constructs. By contrast, in the non-integration task, only time to reach the last page correlates with task time significantly. This confirms the previous analyses; people with the integration task continued to work on integration after they finished going through information sources; and both time phases contributed to the total time spent on the entire problem solving. However, the total time spent on the non-integration task can be represented primarily by time to reach the last page. These results suggest that the two tasks induced different kinds of processing.

It is also notable that task times in both types of task show marginally significant correlations with how frequently people went back to the rules. The more often participants went back to the previous pages and referred to the rules, the longer time they became to consume in total. However, at a more fine-grained level, the tasks were quite different here. Particularly, the significantly high positive correlation between time to reach the last page and rule counts only for the non-integration task shows that the task time in the non-integration task heavily originated from the time spent going back to the rule pages. By contrast, time to reach the last page did not correlate with rule counts in the integration task suggesting that the layout effect was not just an extra page flipping effect.

As for the unexpected pattern of free recall, this correlation table renders a clue. The relationship between free recall and the final integration time shows a significant negative correlation, which implies that as people spent extra time on the final integration phase, they might have lost their opportunity to benefit from a recency effect. However, it still leaves unexplained why participants could remember more information on the 17th page out of 20 pages. It does suggest a better design would be to interrupt participants before they go into the final integration phase and give them a recall test prior to final integration. This method would clarify further the unsettled issues among spatial cueing and strategy hypotheses.

Table 4. Mean, SD, Cohen's d for the spatial layout effect and, Intercorrelations Between Variables for Integration Task and Non-integration Task

	M	*SD*	*d*	1	2	2-1	2-2	3	4	5
Integration Task (n = 39)										
1. Practice time	12.5	2.0	0.7	–	.40*	.02	.08	.09	-.42*	.25
2. Task time	14.0	4.7	1.8		–	.62*	.89*	.07	-.39*	.43†
2-1. Time to reach the last page (n = 20)	7.7	1.8				–	.18	.05	-.09	.03
2-2. Final integration time (n = 20)	7.5	3.1					–	-.05	-.53*	.25
3. Task accuracy	7.7	1.5	0.2					–	-.16	.27
4. Free recall (n = 34)	2.4	1.1	1.0						–	-.27
5. Rule count (n = 16)	9.8	8.5	1.4							–
Non-integration Task (n = 39)										
1. Practice time	4.7	.72	.42	–	.14	.37	-.37	.01	.06	.27
2. Task time	8.3	1.5	.01		–	.9*	-.02	-.08	-.09	.54†
2-1. Time to reach the last page (n = 10)	6.5	1.5				–	-.45	-.31	.28	.74*
2-2. Final integration time (n = 10)	1.4	.67					–	.25	-.12	-.58
3. Task accuracy	27.6	1.3	.19					–	-.05	-.20
4. Free recall	1.4	1.4	.44						–	.14
5. Rule count (n = 10)	13.3	3.5	1.3							–

Note. *p < .05. †p < .1.

4.0 GENERAL DISCUSSION

Both studies consistently showed a time benefit effect from the side-by-side layout of information. In study 1, participants who had information pages in the side-by-side format solved an integrative task (i.e. the deer task) 32% faster than those who worked with the overlaid format without any difference in accuracy. Besides this replicated time benefit effect, study 2 demonstrated that the time effect of the side-by-side format belongs to an integrative task, with no effect at all on a non-integrative task such as simple arithmetic problem solving. It also suggested that integrative tasks induce different cognitive processes from non-integrative tasks in terms of memory and task progress. While subjects who solved a non-integrative task showed a typical recency effect in forgetting function, those who solved an integrative task showed a reversed-recency effect; more items in the 4^{th} to the last page were remembered than from any other more recent pages and the side-by-side format group was better. Moreover, people who solved an integrative task showed a curious pattern of task progress; the side-by-side format group spent far less time in the final integration phase and they maintained the overall fast speed either by arriving the last page faster than the overlaid format group or by finishing up the final integration faster when they used similar amount of time on reaching the last page to the overlaid group. As a whole, these results support the idea that the side-by-side layout of information is a fairly time-efficient format for integrative tasks.

The tentative usefulness of the side-by-side format was also suggested in past reading and writing research. O'Hara and Sellen (1997:[16]) compared reading from paper to reading on-line and found that paper has merits over on-line reading in aspects of making annotative notes, quick navigation, and flexibility of spatial layout. They actually observed that subjects in both conditions preferred to lay out the pages in space when the subjects wanted to grasp the overall structure of the given documents, to refer back other pages, and to plan out an essay. The observed differences between reading paper and on-line reading in O'Hara and Sellen's study (1997:[16]) presumably originated from the nature that paper can be easily spread out in space, not from the medium.

In a more direct manner, Wiley (2001:[24]) and Anderson (2007:[1]) tested advantages of multiple windows browser setting. Wiley (2001:[24]) studied the relationship between browser design and learning gains in essay writing, which found that students showed deeper understanding of the source information in their argument essay when they read the source documents in a two side-by-side windows browser design. However, students did prefer a single window browser to a two-window browser setting. This preference resulted in adding a third condition of a two-window browser and specific instructions on why they have to use both windows. More learning gains were observed in both groups using a two-window browser (with or without instructions). Similar to these results, in a recent technical report of a monitor display study done for NEC Display Solutions, Anderson and his colleagues (2007: [1]) reported that 17-inch dual-monitors or one large monitor display over 24-inches is a considerably more productive setting when the user works with multiple sources of information, which requires multiple windows to be opened at the same time. As seen in these studies, physical layout of information does have some effect on performance and in particular, the side-by-side format has some positive effect on tasks of an integrative nature.

This phenomenon is considerably surprising because it violates a common understanding of visual perception: detailed information can only be perceived in the fovea ([13]Hirsch & Curcio, 1989). Finely detailed visual input can be detected within 1° of the center of the fovea, and the acuity diminishes sharply after 5°, the region outside of the parafovea ([7]De Valois & De Valois, 1988). People can process information in detail only one at a time. Therefore, even if two information pages are presented together side-by-side, in theory, one can only give attention to one page at a time, more precisely, only one item at a time. In addition to the issue of visual acuity, the capacity of visual working memory is also limited to only three or four objects ([2]Baddeley, 2003). Thus, even when the two pages are laid out side-by-side, items seen in a page usually slips out of one's visual working memory far before one moves on to the other page. This nature of visual perception makes the two physical formats of information (i.e. the side-by-side and the overlaid format) basically equivalent. Yet, a strong time benefit was observed consistently in the experiments.

Although neither of the initial two hypotheses (i.e. the spatial cueing and the strategy) was successful in explaining the time benefit phenomenon, the measurement devices implemented to differentiate the hypotheses did contribute in ruling out various explanations. First, the fatigue of standing during problem solving in the wall condition did not cause shorter task time. In the study 1, participants in the wall condition had no particular instruction on sitting or standing while they were working on the problem. It naturally resulted that about half of the wall subjects were sitting during the task, and the other half were standing. Since not a single difference was observed between the participants who were sitting and who were standing, the data was treated as a whole and underwent the analysis. Further, the non-integration task in study 2 showed no time effect even though the wall subjects were required to stand.

Secondly, medium in itself was not the cause of the time effect. The time benefit in the side-by-side format implemented in the wall condition was consistently detected to the overlaid format regardless of whether it was realized in the computer or in the paper within a folder.

Finally, the memory test ruled out the possibility that backward masking occurs only in the overlaid condition. Because participants remembered significantly more items from the 4th to the last page, there was apparently no evidence for backward masking. In fact, it was exactly the opposite of the theory of backward masking; the later visual stimulus did not erase the previous items in memory and this was true to both conditions.

Another theoretical construct of potential relevance to this phenomenon is situation awareness (SA). In SA, as its definition manifests: “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” ([8]Endsley, 1996), gathering information from multiple sources is often considered as an important task in such domain that users are dealing with complex systems (e.g. aviation, air traffic control, navigation, and nuclear power), In fact, one of the goals of SA measurement tools is to “provide an index of how well operators are able to acquire and integrate information in a complex environment where a lot of data may vie for their attention” ([9]Endsley, 1998). Moreover, it is reported that over 35% of operational errors were occurred in enroute Air Traffic Control operations because of failures to correctly perceive, integrate or comprehend information, which includes failure to monitor data, memory loss, and lack of or poor mental model ([10]Endsley, 1999). The time benefit showed from this research can offer an implication to the display designs in such systems in the way that the side-by-side format can promote rapid detection, reduce memory loss, or mitigate time pressure.

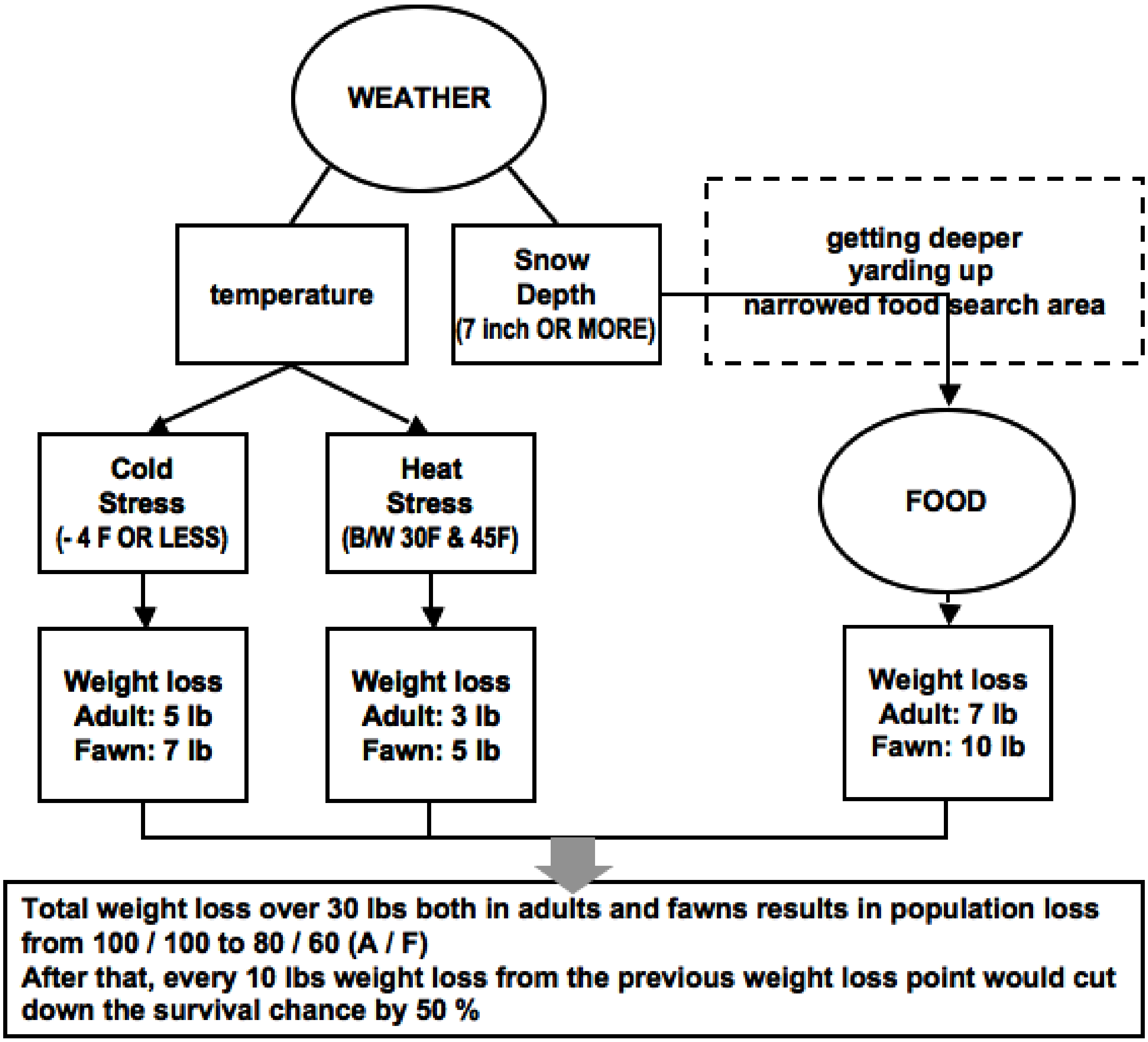
The advantages of the side-by-side format can be applied not only to time constrained jobs but also to productivity in most jobs in general. Given that the results from the deer task studies showed a strong advantage of the side-by-side format, promoting multiple and larger display settings could be a convincing suggestion. As Anderson (2007:[1]) observed, dual-monitor or a large monitor workstation might yield higher productivity than a typical single monitor set up does by enabling the users to work with multiple side-by-side windows when it is needed.

However, there are remaining questions to be answered for full implementation. First of all, it is unclear whether the time effect is due to some kind of memory benefit. Although the forgetting functions implicated partial supports for a memory benefit existing in the side-by-side format, the timing of the memory test has entailed a problem. Participants took the memory task at the end of the entire task, so the time lapse between when they reached the last page and when they received the test fluctuated by subjects. Some subjects, especially many of the paging subjects, spent longer time in final integration phase, which consequently made the information they saw became far older comparing to other fast subjects. To fix this problem, a follow up study with a memory test given before final integration phase is in progress. While the modified follow up study is expected to remove a confounded factor, another study would be worthwhile to conduct to test differences purely in a memory-wise manner. A memory task situated in two different formats could directly tell whether the side-by-side format assists limited memory.

Besides the memory issues, there is a human factor question; why people usually prefers a single window setting? In both of Wiley (2001:[24]) and Anderson (2007:[1])'s studies, people initially preferred single monitor set up and refused multiple window browser or dual monitor setting when they were allowed to choose. Moreover, a preliminary study done with NASA also finds this tendency. In situ observation coding showed that most scientists typically used laptop (i.e. single small monitor) during the mission, even though they were doing integrative tasks most of the time. Since implementing theoretical findings often faces unexpected resistance, this seemingly odd preference for a single monitor would be worth to look into.

Lastly but not the least, expertise can be an important factor to consider. The meteorologists, who offered the starting point of this research, were intermediate weather forecasters, and students participated in this research were all novice in terms of the specific task (i.e. the deer task and the symbol task). Thus, no study is available that foresees how experts in a domain would deal with the information in the two different formats. To effectively apply these findings in real world, a study in association with expertise is indispensable.

APPENDIX A. Summary; A diagram of rules for prediction



APPENDIX B.

		Disagree				Agree
1	Before the task started, I thought to write down the rules in my notes.	1	2	3	4	5
	During problem solving,					
2	I had trouble remembering information from prior pages and often had to look back.	1	2	3	4	5
3	I generally relied on information being available on paper instead of trying to remember it.	1	2	3	4	5
4	I had trouble finding the information I need.	1	2	3	4	5
5	All the information from different pages blurred together.	1	2	3	4	5
6	I had trouble remembering the rules.	1	2	3	4	5
7	I purposely tried to keep temperatures and snow amounts in my head.	1	2	3	4	5
8	I got information from one page confused with information on another page.	1	2	3	4	5
9	I got lost about where I was in the task (e.g. what page).	1	2	3	4	5
10	I found the task interesting.	1	2	3	4	5
11	I was trying to get it all done correctly.	1	2	3	4	5
12	I was trying to get it all done quickly.	1	2	3	4	5

BIBLIOGRAPHY

[1] Anderson, J. A., Hill, J., Parkin, P., and Garrison, A. (2007). *Productivity, screens, and aspect ratios; A comparison of single, traditional aspect, dual, traditional aspect and single, widescreen aspect computer displays over simulated office tasks across performance and usability.* (CIC Report No. 200719: NEC Display Solutions). Illinois.

[2] Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Review: Neuroscience, 4,* 829-839.

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

[4] Blessing, S. & Anderson, J.R. (1996). How people learn to skip steps. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 576-598.

[5] Card, S.K., Mackinlay, J.D., & Shneiderman, B. (1999). Readings in Information Visualization: Using Vision To Think. Morgan Kaufmann Publishers, Los Altos, CA.

[6] Chen, C. (1999). Information Visualization and Virtual Environments. Springer, London.

[7] De Valois, R. L. & De Valois, K. K. (1988). Spatial vision. New York: Oxford University Press.

[8] Endsely, M. R. (1996). Situation Awareness Measurement in Test and Evaluation. In T. G. O'Brien and S. G. Charlton (Eds.), *Handbook of Human Factors Testing & Evaluation* (pp.159-180). Mahwah, NJ: LEA

[9] Endsely, M. R. (1998). *A comparative analysis of SAGAT and SART for evaluations of situation awareness*. Paper presented at the 42^{nd} Annual Meeting of the Human Factors & Ergonomics Society, Chicago, IL.

[10] Endsely, M. R. (1999). Situation awareness and human error: Designing to support human performance. *Proceedings of the High Consequence Systems Surety Conference.* Albuquerque, NM.

[11] Fu, W. T. & Gray, W. D. (2000). Memory versus perceptual motor tradeoffs in a Blocks World task. *Proceedings of the Twenty-second Annual conference of the Cognitive Science Society* (pp154-159) Hillsdale, NJ: Erlbaum.

[12] Gray, W. D., Schoelles, M. J., & Sims, C. R. (2004). Learning to choose the most effective strategy: Explorations in expected value. *Proceedings of the sixth International Conference on Cognitive Modeling* (pp. 112-117). Pittsburgh, PA: Carnegie Mellon University/University of Pittsburgh.

[13] Hirsch, J. and Curcio, C.A. (1989). The spatial resolution capacity of human foveal retina. *Vision Research. 29,* 1095-1101.

[14] Larkin, J. and Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science. 11*, 65-99.

[15] Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory & Cognition. 28,* 411–421.

[16] O'Hara, Kenton and Sellen, Abigail. "A Comparison of Reading Paper and On-Line Documents", *Proceedings of CHI '97, Human Factors in Computing Systems* (pp 22-27). Atlanta, Georgia.

[17] Rothkopf, E. Z. (1971) Incidental memory for location of information in text. *Journal of Verbal. Learning and Verbal Behavior, 10,* 608-613.

[18] Shah, P. & Hoeffner, J. (2002). Review of graph comprehension research: implications for instruction. *Educational Psychology Review, 1*, 47-69.

[19] Trafton, G., Trickett, S., Schunn, C. & Kirschenbaum, S. (2007). *Complex visualizations, spatial transformations, and uncertainty.* Presented at the meeting of the Naval Research

Laboratory, Washington, D.C.

[20] Tufte, E.R. (2001). The Visual Display of Quantitative Information. Graphics Press, Chesire, CT.

[21] Vessey, I. (1991). Cognitive fit: a theory-based analysis of the graphs versus tables literature. *Decision Sciences. 22*, 219–240.

[22] Vincow, M. A. & Wickens, S. D. (2007). Spatial layout of displayed information: Three steps toward developing quantitative models. *Human Factors and Ergonomics Society Annual Meeting Proceedings, Computer Systems. 5*, 348-352.

[23] Ware, C. (2000). Information Visualization: Perception for Design. Academic Press, San Diego, CA.

[24] Wiley, J. (2001). Supporting understanding through task and browser design. *Proceedings of the Twenty-third annual Conference of the Cognitive Science Society,* (pp. 1136-1143). Hillsdale, NJ: Erlbaum.