(T&F) TBIT100403

BEHAVIOUR & INFORMATION TECHNOLOGY, 2003, VOL. 0, NO. 0, 00–00

Taylor & Francis
Taylor & Francis Group

Layout attributes and recall

MARKETTA NIEMELÄ and PERTTI SAARILUOMA

Department of Computer Science and Information Systems, University of Jyväskylä, P.O.Box 35, 40014 Jyväskylä, Finland; e-mail: niemark@cc.jyu.fi and e-mail: psa@cc.jyu.fi

Abstract. The spatial arrangement of elements such as icons in a computer interface may influence learning the interface. However, the effects of layout organization on users' information processing is relatively little studied so far. The three experiments of this paper examined two attributes of layouts: spatial grouping by proximity and semantic coherence. Learning was assessed by tasks in which 30 participants recalled icon-like items' labels, locations, or both as a series of study-recall trials. The results show that layout organization interacts with task demands. Semantic organization improves recall of labels, and spatial grouping supports recall of locations. When both labels and locations are learned concurrently, the best recall performance is associated with a simultaneously grouped and semantically coherent layout. However, semantic and spatial organization may interact unexpectedly on learning. The findings are discussed from the viewpoint of information chunking in memory processes and interface design.

1. Introduction

An important role for human-computer interaction research is to gradually transform interaction design from intuitive to explanatory. This means that design solutions are increasingly often based on empirically supportable psychological principles instead of intuitive ideas about ideal interfaces. Introspective and subjective views are necessarily narrow and thus design risks are greater in intuitive approaches. However, before full-scale explanatory design is possible we must have a clear understanding about how such basic psychological principles as chunking affects on various stages of interaction. This is a non-trivial issue, because the same principles may flesh out very differently in different situations and we have to individually investigate the individual cases.

We know, for example, that people have to learn screen layouts to smoothly interact with them. However, this learning process is a complicated scientific process as interfaces have a manifold of dimensions. We should understand how users integrate information about spatial locations, and we should, equally importantly, have a clear idea about the roles of various semantic elements such as textual labels in learning to use an interface. In addition, we should have an idea about how various types of information on a screen are integrated in mental representations. The only way to find answers to these important prerequisites of explanatory design is patient empirical work.

In interacting with screen layouts, spatial grouping is vital and recommended in many guidelines of interface design (Gittins 1986, Marcus 1995, Shneiderman 1997). When interacting with a computer system, the user forms a mental representation of the interface, which entails such information as the elements, their semantic types and location. Each of these attributes has some effect on the interaction with the interface.

A subtask in interaction is visual search for screen elements such as buttons, labels, icons, and menu elements. The visual appearance of the elements influence their recognizability and thereby the speed and ease of their use. For instance, visually complex icons are more difficult to search than simple icons, which are composed of few visual features (Byrne 1993). Therefore, if the icon pictures are abstract, meaningless, or not clearly associable to their meaning, users then will rely on their memory for location rather than the appearance of the icon during search (Blankenberger and Hahn 1991, Moyes 1994, Ehret 2002).

Spatial attributes other than positional constancy (e.g. Teitelbaum and Granda 1983, Green and Barnard 1990) of interface elements are not much studied. For instance, studies concerning the use of icons can be generally divided into two groups, those focusing to the appearance of icons, and those comparing icons and text in learning and performance (Wiedenbeck 1999). It is very difficult to find studies about icon layout. This is interesting considering that according to Gittins (1986),

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0144929031000156924

one of the reasons for icons' efficiency may be that they encourage the user to explore the visual relationship and organization of objects more than with verbal commands. Element grouping in the interface may thus influence learning and recollection in the operation of computer-based systems. Likewise, Tullis (1997) suggests that organization of interface elements influences the user's ability to extract information on the display and interpret it.

In this paper, we shall systematically investigate two organizational attributes of items, spatial grouping and semantic coherence. These attributes are illustrated in figure 1. Spatial grouping means perceptual grouping by proximity of items (figure 1B and D). Semantic coherence refers to positioning of items so that semantically related items are close to each other (figure 1C and D). Semantic coherence does not necessarily assume spatial grouping in the layout. Items can be arranged semantically coherently without forming spatially distinct groups.

Most HCI guidelines include instructions for how to design the interface layout properly, following the Gestalt laws of perceptual organization of visual information (Wertheimer 1923). Humans tend to form perceptual groups of elements, which are close or similar to each other. Spatial grouping, or grouping by proximity, is a basic organization principle of items in visual displays. Although spatial grouping is not the strongest of the perceptual grouping principles (Palmer 1992), it is among the fastest. For example, it is faster perceived than grouping by similarity of items (Ben-Av and Sagi 1995, Han *et al.* 1999).

Good grouping of interface elements based on these laws helps the user to effortlessly perceive the structure of the interface. In Marcus' (1995) guidelines for user interface design, organization is one of the main principles in design. In screen layout, this means using a grid structure, standardizing the screen layout, and grouping related items (and dissociating unrelated items). Gittins (1986) advises to group function icons by their logical associations. The user is then able to process the functions as cognitive chunks, and perform faster search on the icons.

More specific guidelines are provided by Tullis (1997). On the basis of his experiments, he recommends arranging interface objects by their semantic relationships into small (less than 5 degrees in diameter) spatial groups. A small group needs to be fixated only once for information extraction, whilst a large group requires more fixations. The spatial and semantic aspects of the layout were not tested separately, as Tullis (1997) states that the user should be able to assume that the elements within a group relate to each other semantically.

Other grouping studies concern the layout of menu items and icons in the interface. Categorical organization in menus enabled faster search for target menu items than random arrangement (Card 1981, McDonald *et al.* 1983, Halgren and Cooke 1993). Also alphabetical ordering is faster to search than random one, but categorical organization may be more beneficial because it can have longer-term influences as the user can develop a more appropriate conceptual model of the system via the organization of information in menus (Halgren and Cooke 1993).

Systematically grouped icons are searched faster than icons positioned randomly on the screen (Niemelä and Saarinen 2000). The icons in the study belonged to four visual groups, thus they could be grouped by their visual similarity independent of the spatial organization, which obscures the effect of mere spatial grouping.

Spatial grouping and semantic coherence are often confounded in these guidelines and studies. In addition, these studies concern mostly search in the interface, not recall. One recall study in which these attributes are treated as separate is reported by Bentum (1998). In the study, levels of conceptual relatedness (semantic coherence) and spatial clustering of a layout of 30 icons were varied. The participants performed better in short-term location retrieval when the layout was conceptually related. This paper aims to propose more deep knowledge of how spatial grouping and semantic coherence in a layout affect the user's interface learning by focusing

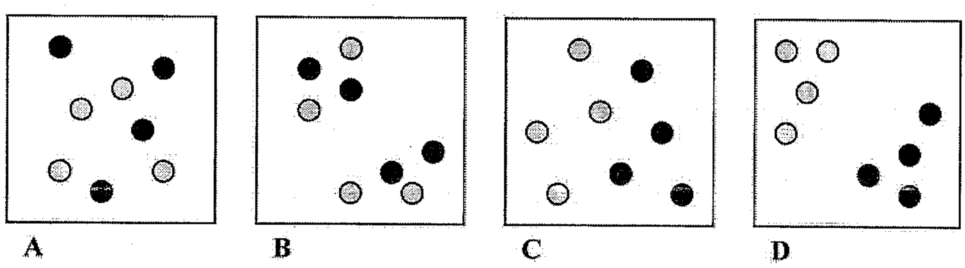

Figure 1. Examples of spatial grouping and semantic coherence. Black and gray items denote different semantic categories. (A) Spatially ungrouped, semantically incoherent layout. (B) Spatially grouped, semantically incoherent layout. (C) Spatially ungrouped, semantically coherent layout. (D) Spatially grouped, semantically coherent layout.

on both location and label recall. To form specific hypotheses for our experiment, we first review studies from the cognitive psychology domain.

The meaning of semantically coherent organization in verbal recall is well established. In studies on free-recall of word lists, words from the same semantic category are usually recalled together, even when their presentation order is randomized. If related words are positioned adjacent to each other, recall is enhanced (Glanzer 1969, Neely and Balota 1981, Toglia *et al.* 1997, Howard and Kahana 2002). In addition, when words are not in a list but spatially distributed, the spatial proximity of associated words facilitates their encoding as semantic chunks (McNamara and LeSueur 1989), and thus improves recall. Based on this, semantically coherent layout should improve label recall.

There is also some evidence of mere spatial grouping affecting recall of verbal material. Words grouped in three columns were recalled slightly better than words in a single column (Decker and Wheatley 1982). This was assumed partly to be due to more distinct locations of grouped words. In another study, a series of eight digits was recalled better if it was spatially broken down into three or four parts instead of one sequence (Magnussen *et al.* 1997). This also was explained in terms of better distinctiveness of grouped digits compared to a uniformly spaced series. Both of these studies used material that is not semantically related, thus they leave open the question whether spatial grouping improves recall of semantically categorized labels. We hypothesize that there is some facilitation of spatial grouping due to distinctiveness of groups enabling better memorability, although the effect of semantic coherence is much stronger.

During interaction with an interface, the user forms an internal spatial map of interface elements (Teitelbaum and Granda 1983, Ehret 2002). Semantically coherent layout may affect location recall by facilitating mental 'chunking' of the elements. Even spatial organization is important in the construction of chunks (Saariluoma and Sajaniemi 1989, 1991). McNamara and LeSueur (1989) studied spatial encoding between item pairs, and found that spatial relations between the verbal items were more likely to be encoded if the items were also semantically related. Correspondingly, expert participants who were able to chunk functional units in symbolic drawings (circuit diagrams), performed better in later reconstruction of the drawings than novices (Egan and Schwartz 1979).

Spatial grouping of elements may influence location recall by bringing visual regularity to the display and reducing visual complexity. If items are positioned in a matrix instead of random positioning, the display is less visually complex. This improves short-term recall of the item locations (Kemps 1999). In addition, simpler visual patterns are easier to reproduce from memory than more complex ones, but it is possible to decrease memory load by chunking (Attneave 1955, Saariluoma 1994). Thus, we expect spatial grouping entail better location recall of items compared to random positioning.

Based on the studies presented above, our hypotheses are proposed as follows:

- Both spatial grouping and semantic coherence improve label recall (Experiment 1).
- Both spatial grouping and semantic coherence improve location recall (Experiment 2).
- Semantic coherence improves recall more than spatial grouping (Experiment 3).

The task in the first experiment is label recall without any spatial memory requirements, and correspondingly, the task in the second experiment focuses on spatial recall excluding the verbal factors. In the third experiment, the tasks of label and location recall are combined, which enables us to directly compare the layout organization effects on the two recall tasks. In addition, there is a new dimension, namely label-to-location assignment. This task is more difficult than mere position reconstruction and is partly verbal in nature (Postma and De Haan 1996). Thus, we can expect that semantic coherence is a stronger factor in recall in Experiment 3 than spatial grouping.

2. General method

In all three recall experiments, the apparatus, stimuli, and partly the procedure were the same. The experiments were implemented on a personal computer using the Microsoft® Windows®2000 operating system and a 19-inch sized display. The experiment consisted of four layout conditions, each differentiated by spatial grouping and semantic coherence of a display of icon-like items attached with textual labels. In the Ungrouped Incoherent condition, the items were randomly positioned on the display (as in figure 1A). In the Grouped Incoherent condition, the items were positioned in three spatial groups of four items each, in an unsystematic way (see figure 1B). In the Ungrouped Coherent condition, the items from the same semantic category were placed close to each other, but so that no spatially explicit groups were formed (the items were uniformly distributed) (see figure 1C). In the Grouped Coherent condition, the items were spatially grouped by their semantic category (see figure 1D).

The test stimuli were a display of 12 labelled items. An item was a green square (figure 2), sized 0.34° × 0.34° viewed from a distance of 50 cm. All the items were similar in appearance, but they had individual labels positioned 0.17° below the square. The height of the letters (10-point Courier New) was 0.11°. The labels used came from 12 semantic categories, and a different set of three categories was used in each layout condition. The 12 semantic categories are listed in the Appendix. The labels were 4–9 letters long, and approximately equal in frequency.

Ten participants were recruited for each experiment for a small monetary reward. The participants were all university students and were tested individually. In each experiment, the participant's task was to study the item display and then perform a recall task immediately, as a series of study-recall trials. The task was different in each experiment, requiring either memory for labels (Experiment 1), locations (Experiment 2), or both labels and their locations (Experiment 3).

The participants had 20 s time to study the display and then they performed the recall task. The time for recall was not limited. The participant finished the trial by pressing a key in the keyboard. If recall was not faultless, the same stimulus display was shown again for re-studying. The task was finished either when the participant succeeded completely in recall or the limit of 10 trials was reached (figure 3). No performance feedback was provided to the participant. The task was repeated four times, once in each layout condition (within-subjects design). The labels were different in the four conditions, and the same label set was used in all three experiments. The order of the layout conditions was randomized for each participant. Before each task, a short practice session was conducted, consisting of three items from two categories (fabrics and a building). The items were grouped according to the layout condition. The participant was shown the practice stimulus display for five seconds in order to study the labels of the items. After that, the recall was performed similarly to the actual task.

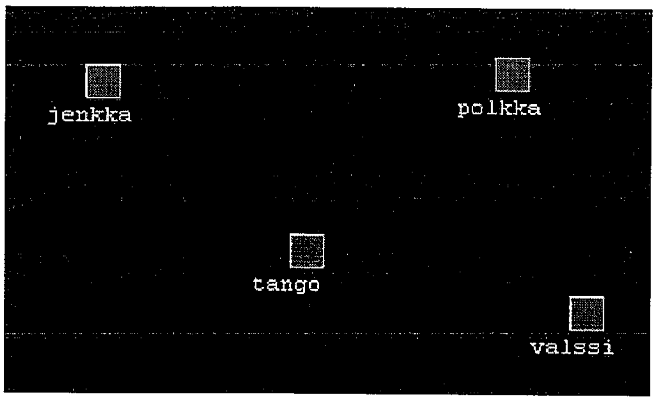

Figure 2. Items from a semantic category 'Dances'.

3. Experiment 1

In Experiment 1, participants were asked to memorize the labels of the items positioned on the display. This task did not require any use of spatial information during learning or recalling. The question of interest was whether the participants are able to benefit from spatial grouping in addition to semantic coherence in this verbally directed task.

3.1. *Method*

The 10 participants were aged 19–29 years, eight of them were female and two male. In the test situation, the participant was shown the stimulus display for studying the labels. The items were positioned in the display according to one of the four layout conditions. After 20 s studying and 500 ms pause (black screen), an empty window was shown which allowed the participant to write the labels in a list. The list had a limit of 12 items. The recall was complete when all 12 labels were recalled correctly. Every participant performed four study-and-

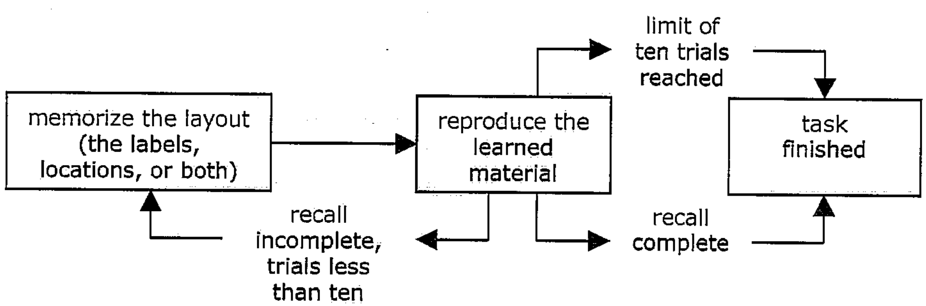

Figure 3. Schematic presentation of a trial in the experiments.

recall tasks, one of each layout condition, in a random order.

3.2. *Results*

The data were examined by analyses of variance (ANOVAs) with spatial grouping and semantic coherence as fixed factors, and participant as a random factor. The dependent variables were task performance, task completing time from those participants who succeeded to complete, and correctness of label recall. 'Task performance' refers to the trial in which the participant completed the task. The trial was scored so that completing the task in the first try was scored as 10, the second try was scored as 9, etc. If the participant did not succeed in the task during the last (tenth) trial, the performance was scored as zero. In this experiment, all 10 subjects were able to complete the recall tasks in all four layout conditions (table 1).

Performance was significantly influenced by semantic coherence. In the semantically coherent layout, the task was completed in an earlier trial on average than in the incoherent layout [$F(1, 9) = 21.0$, $p = 0.001$]. Completion was also faster when the layout was semantically coherent [$F(1, 9) = 10.3$, $p = 0.011$] (table 2). Spatial grouping had no significant effect on performance [$F(1, 9) = 1.7$, $p > 0.1$] or completion time [$F(1, 9) = 0.1$, $p > 0.1$]. There were no significant interaction effects between grouping and coherence.

Figure 4 shows the correctness of label recall in the four tasks by trial. Recall correctness was analysed with a three-way ANOVA, trial as the third fixed factor. Trial had a significant main effect on recall [$F(9, 81) = 16.3$, $p < 0.001$], which confirms learning during the task.

Table 1. The completion trials in Experiment 1, from those subjects who successfully completed their task (their number in parentheses)

Spatial grouping	Semantic coherence: Incoherent	Semantic coherence: Coherent	Mean
Ungrouped	3.4 (10)	1.7 (10)	2.6
Grouped	2.7 (10)	1.6 (10)	2.2
Mean	3.1	1.7	2.4

Table 2. Completion times (s) in Experiment 1

Spatial grouping	Semantic coherence: Incoherent	Semantic coherence: Coherent	Mean
Ungrouped	133	70	101
Grouped	126	68	97
Mean	129	69	99

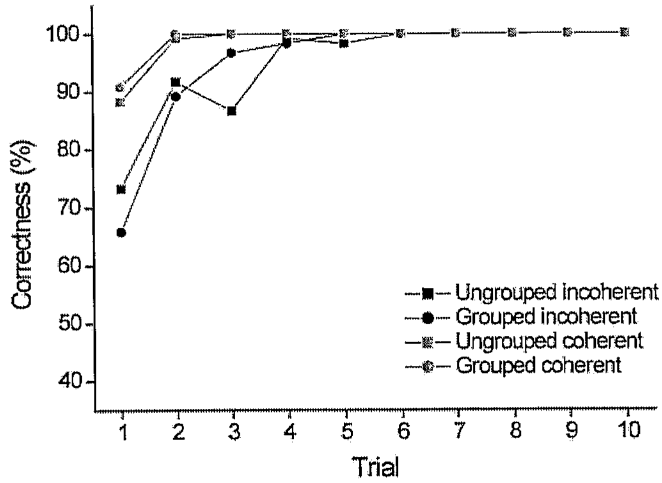

Figure 4. Correctness of label recall by trial in Experiment 1.

Semantic coherence had a main effect on correctness of recall. Recall of labels was more accurate when the layout was semantically coherent [$F(1, 9) = 15.3$, $p = 0.004$]. There was interaction between semantic coherence and trial [$F(9, 81) = 7.9$, $p < 0.001$]. The positive effect of semantic coherence of the layout was greater in the beginning of the task than in the latter trials (because the participants learned all the labels). Spatial grouping had no effect [$F(1, 9) = 0.1$, $p > 0.1$] on correctness, and no other significant interactions were found.

3.3. *Summary and discussion*

As expected, semantic coherence in the layout arrangement had a significant effect on recall of labels. In the semantically coherent layout, the participants learned the labels faster and more effectively compared to the incoherent layout. Recall of labels was more accurate from the beginning of the task in the coherent layout. Contradicting to our hypothesis, spatial grouping in the layout had no significant main or interaction effects on learning of labels. In previous studies, spatial grouping has been found to improve word recall (Decker and Wheatley 1982, Magnussen *et al.* 1997). Possibly the clear categorization of words directed the participants to process the semantic organization so strongly that grouping was insignificant.

However, when grouping was applied with semantic coherence, it should have made the categories more distinctive and thus improve memory. In fact, in the first two trials this was true. The easiness of the recall task may have obscured the underlying effect. In Experiment 3, the recall task is more demanding, requiring memory for both labels and location. The effect of grouping in making semantic categories distinctive and thus easier to memorize may be observable there. Before that, we test location recall in the same layout conditions as in the Experiment 1.

4. Experiment 2

In Experiment 2, we examined how spatial and semantic arrangement of a layout affects location recall. The participants memorized locations of labelled items in the display and in recall, positioned the items according to their spatial memory. The labels were covered, thus this task required mere reconstruction of locations.

4.1. *Method*

Ten new participants, eight female and two male, were aged 20–25 years. The experimental display of 12 items each was the same as in Experiment 1. After a study phase, a display was shown in which 12 items labelled with '*****' were positioned in a row in the bottom of the screen. The participant could restore the original display by mouse-dragging the items one by one to their locations.

The location of an item was defined as correct if the item was placed within a circular area around a location, which was previously occupied by any item in the study display. The radius of the area was 120 pixels, which was six times the length of an item side.

4.2. *Results*

The data analyses with ANOVAs were similar to the analysis in Experiment 1, except the correctness measures were about locations of the items. Task performance was scored similarly to Experiment 1.

The completion trials and times of those participants who successfully completed the tasks are shown in tables 3 and 4. Spatial grouping of the layout improved performance [$F(1, 9) = 6.1$, $p = 0.036$]. Semantic coherence did not affect performance [$F(1, 9) = 0.9$, $p > 0.1$], and there was no interaction effect between these two factors [$F(1, 9) = 0.6$, $p > 0.1$].

Coherence or grouping had no main effects on completion time [$F(1, 9) = 0.1$, $p > 0.1$, $F(1, 9) = 3.0$, $p > 0.1$, respectively] but they interacted with each other [$F(1, 9) = 9.9$, $p = 0.017$]. Spatial grouping improved location learning only when implemented in the semantically incoherent layout [$F(1, 9) = 8.7$, $p = 0.016$]. Similar analysis to performance showed that the positive effect of spatial grouping was significant only when the layout was incoherent [$F(1, 9) = 7.0$, $p = 0.027$].

Similar to Experiment 1, trial was added as a third fixed factor to the analysis of correctness of location recall (figure 5). Correctness was significantly dependent only on trial, which confirms learning during the task [$F(9, 81) = 10.0$, $p < 0.001$]. Spatial grouping approached significance [$F(1, 9) = 3.9$, $p = 0.079$], which is in line with the above results of grouping supporting performance. Semantic coherence had no effect [$F(1, 9) = 0.5$, $p > 0.1$]. However, when comparing semanti-

Table 3. The completion trials in Experiment 2, from those subjects who successfully completed their task (their number in parentheses)

Spatial grouping	Semantic coherence: Incoherent	Coherent	Mean
Ungrouped	3.1 (10)	2.0 (9)	2.6
Grouped	1.3 (10)	2.0 (9)	1.7
Mean	2.2	2.0	2.1

Table 4. Completion times(s) in Experiment 2

Spatial grouping	Semantic coherence: Incoherent	Coherent	Mean
Ungrouped	217	125	174
Grouped	63	139	101
Mean	140	133	136

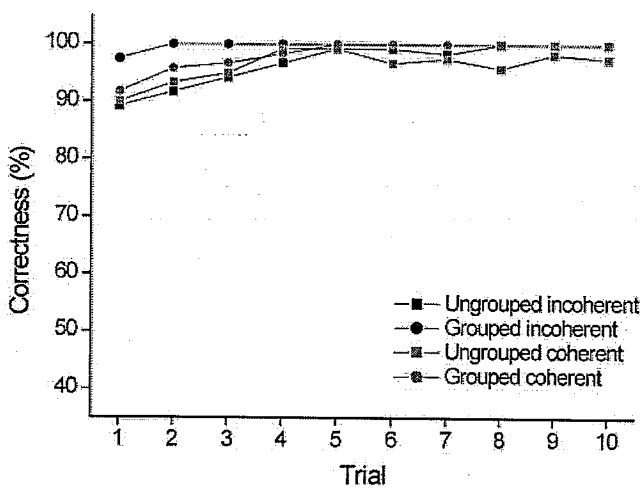

Figure 5. Correctness of location recall by trial in Experiment 2.

cally coherent and incoherent layouts, it was found that the facilitative effect of grouping on recall correctness was significant when the layout was also semantically incoherent [$F(1, 9) = 7.4$, $p = 0.023$].

All three factors – trial, semantic coherence, and spatial grouping – interacted with each other [$F(9, 81) = 3.5$, $p = 0.001$]. Particularly in the beginning of the task, grouping facilitated learning (mostly when implemented in semantically incoherent layout). In the latter trials, recall in ungrouped layout improved to the level of grouped layout.

4.3. *Summary and discussion*

Spatial grouping of layout had a significant effect on performance, as the recall task was, on average, completed earlier than with the ungrouped layout. In line with this, the positive effect of spatial grouping on accuracy of location recall also approached significance.

Semantic coherence of the layout had no significant effects on recall. However, it interacted with spatial grouping. Spatial grouping improved recall mostly when the layout was semantically incoherent. This pattern was observable in all measures of recall, and it is not easy to explain. Possibly semantic coherence interferes with the positive influence of spatial grouping, maybe by pulling the participant's attention too much to the categorical organization of the items during the study phase. The participants turn to encode locations of clusters instead of single elements' locations. Experimental support for this explanation is provided by Hirtle and Mascolo (1986). They showed participants spatial arrays of place names from two semantic clusters (e.g., Playground from the recreational facilities cluster and Bank from the city buildings group). The names were arranged semantically coherently in the map according to their clusters. The participants' recall of locations was affected by the semantic clustering so that an item's location was biased closer to the locations of the other items in the appropriate semantic cluster. Thus, the recall of locations may have suffered in accuracy because of the strong semantic organization in the spatially grouped and semantically coherent layout.

The last experiment allowed us to directly compare the effects of layout organization on label recall and location recall.

5. Experiment 3

In Experiment 3, we were interested in the interaction of verbal items and their locations in learning. The previous experiments showed that semantically coherent organization of the items helps in label memorization (Experiment 1), and spatial grouping of the items benefits memory regarding locations (Experiment 2). The latter experiment also indicated an interesting interaction between coherence and grouping. Here we studied how these factors influence during a task that requires memorization for both labels and locations, and assigning labels to their correct locations.

5.1. *Method*

Another 10 participants were aged 20–29 years, eight of them were female and two male. The display of 12 items and the layout conditions were the same as in previous experiments. After a study phase, the participant restored the original display by clicking the mouse in remembered locations. A click made a non-labelled item appear in the location. The participant could enter a label to a specified label area of the item. A non-labelled item could be moved by clicking the mouse in a new location, which made the item jump to that location. No more than 12 items could be positioned to the screen.

Complete recall for finishing the task was defined by means of recalling both the locations and the labels of the items correctly. The criterion to location correctness was the same as in Experiment 2. In addition, labels had to be recalled completely, and they had to be assigned to the same locations as in the study display.

5.2. *Results*

The data was analysed with ANOVAs, similar to the previous experiments. The dependent variables were task performance, task completion time, and correctness of recall of labels, locations, and label-location assignments. Task performance was scored as in Experiment 1.

The recall task in this experiment was a difficult one, as the task was completed in the last rather than in the first trials, and depending on the layout condition, only six to nine participants of 10 were able to complete the task overall (table 5). No significant main or interaction effects of semantic coherence or spatial grouping on performance were found.

The completion time of those participants who completed the task depended on spatial grouping [$F(1, 7) = 5.5$, $p = 0.050$] so that in the grouped layout more time was spent in completing the task (table 6). There were no significant semantic coherence or interactive effects on completion time [both Fs < 1.1, ps > 0.1].

For correctness analyses, trial was included in the ANOVA as the third fixed factor. Correctness of label recall (figure 6) depended on trial [$F(9, 81) = 30.8$, $p < 0.001$], indicating learning during the task. Correctness was enhanced in the semantically coherent layout [$F(1, 9) = 5.5$, $p = 0.043$], whilst spatial grouping had no effect [$F(1, 9) = 1.6$, $p > 0.1$]. Interaction between semantic coherence, spatial grouping, and trial approached significance [$F(9, 81) = 1.9$, $p = 0.061$]. There was a similar interaction with correctness of location recall and assignment as well, therefore they all will be considered collectively.

Correctness of location learning (figure 7) was affected significantly only by trial [$F(9, 81) = 43.3$, $p < 0.001$], not by semantic coherence [$F(1, 9) = 3.2$, $p > 0.1$] or spatial grouping [$F(1, 9) = 0.1$, $p > 0.1$]. Similarly, correctness of label-location assignments (figure 8) was also significantly affected only by trial [$F(9, 81) = 92.9$, $p < 0.001$], not by semantic coherence [$F(1, 9) = 3.1$, $p > 0.1$] or spatial grouping [$F(1, 9) = 0.0$, $p > 0.1$]. For label-location assignment, semantic coherence interacted with trial [$F(9, 81) = 2.3$, $p = 0.024$]. Coherence facilitated assignment in the first trials but not in the latter trials, when recall was close to complete in all layouts.

Interaction between semantic coherence, spatial grouping, and trial was significant also for location recall [$F(9, 81) = 43.3$, $p < 0.001$] and label-location assignment [$F(9, 81) = 2.3$, $p = 0.027$] in addition to label recall. Comparing figures 6, 7, and 8, it seems that

Table 5. Completion trials in Experiment 3, from those subjects successfully completed their task (their number in parentheses)

Spatial grouping	Semantic coherence: Incoherent	Semantic coherence: Coherent	Mean
Ungrouped	5.2 (6)	4.9 (9)	5.0
Grouped	6.1 (7)	4.8 (8)	5.4
Mean	5.7	4.8	5.2

Table 6. Completion times (s) in Experiment 3

Spatial grouping	Semantic coherence: Incoherent	Semantic coherence: Coherent	Mean
Ungrouped	417	337	369
Grouped	587	423	499
Mean	508	377	434

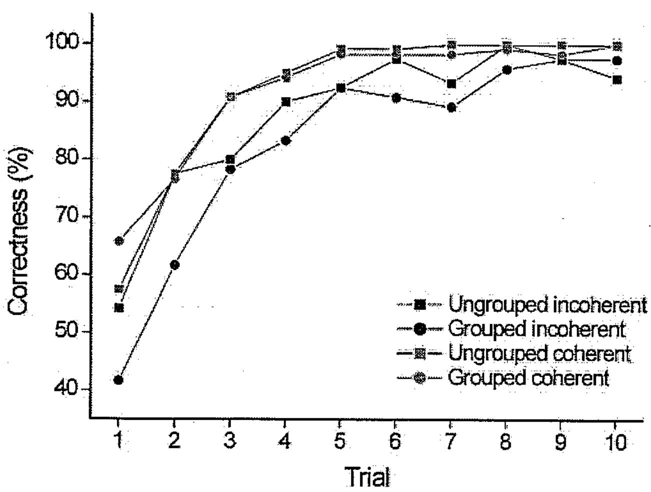

Figure 6. Correctness of label recall by trial in Experiment 3.

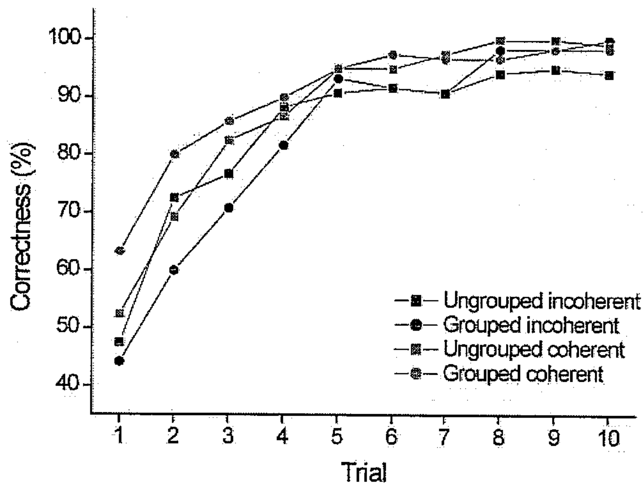

Figure 7. Correctness of location recall by trial in Experiment 3.

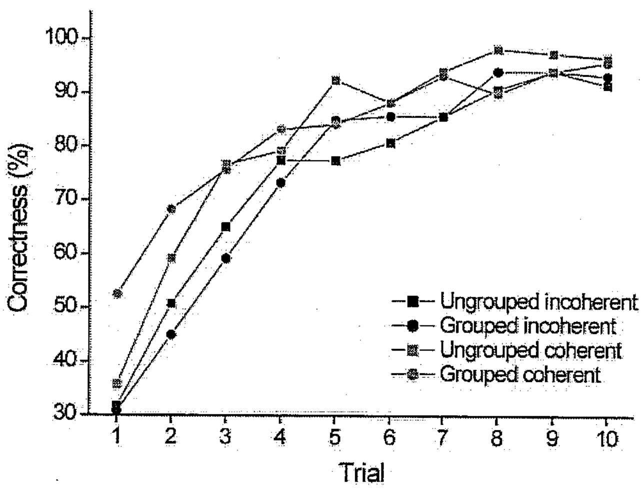

Figure 8. Correctness of label-location assignment by trial in Experiment 3.

the pattern is quite the same in all cases. It seems that semantic coherence facilitated recall significantly only when implemented in parallel with spatial grouping, and especially during the initial learning. Spatial grouping in a semantically incoherent layout rather deteriorated recall. A more detailed analysis confirmed an interaction effect between semantic coherence and trial in spatially grouped layout for all correctness variables [all $Fs(9, 81) > 2.8$, all $ps < 0.007$].

5.3. *Summary and discussion*

In line with the results of Experiment 1, participants learned labels more effectively in terms of recall correctness in the semantically coherent layouts than in the incoherent layout. Locations and their assignment to the labels were learned during the task, but neither spatial nor semantic organization of the layout affected accuracy of their recall.

Spatial grouping influenced completion time, but not as expected. Grouping increased completion time. Performance scores are in the same direction (in favour of the ungrouped layout), although not statistically significant. Deterioration due to grouping seems to be associated especially with semantically incoherent layout. Recall accuracy results support this view. If layout was semantically incoherent, grouping deteriorated also accuracy of recall of labels and locations. On the other hand, grouped and coherent layout was consistently the best layout. These patterns are evident especially in the first trials of the task. In the latter trials, the accuracies become more even. The phenomenon may be due to the false expectations the grouping creates about the semantic coherence in the layout. The same pattern can be found also in the first two trials of Experiment 1.

This experiment was clearly more difficult than the other two. Although the same tasks, label recall and location recall, were present, their combination made this experiment more difficult than experiments 1 and 2. Average accuracy of label recall was 88.6%, which is less than in Experiment 1 (96.9%). In addition, location accuracy of 86.2% was lower than correctness of location recall in Experiment 2 (97.8%). Assigning labels to locations seems to be the most difficult component in this experiment. The average correctness of assignment was 78.4%.

The difference in the difficulties of the experiments is present also in performance levels and completion times (table 7). Both Experiments 1 and 2 had better performance levels than Experiment 3 [$t(51.1) = 8.5$, $p < 0.001$; $t(70.2) = 7.5$, $p < 0.001$, respectively]. Experiments 1 and 3 did not differ from each other [$t(61.6) = 0.1$, $p > 0.1$]. Similarly, completion times were shorter in experiments 1 and 2 than in Experiment 3 [$t(32.1) = 7.1$, $p < 0.001$; $t(41.3) = 5.9$, $p < 0.001$, respectively]. Times in Experiments 1 and 2 were equal [$t(55.8) = 1.6$, $p > 0.1$].

Table 7. Performance and completion times in the three experiments

Experiment	Performance	Completion time(s)
1	8.7	99
2	8.7	136
3	4.4	434
Mean	7.2	205

6. Discussion

The purpose of the three experiments reported here was to enlighten the effects of layout organization on the user's recall of the elements' names and locations. Based on previous research from both the HCI domain and cognitive psychology, we predicted that both spatial grouping by proximity and semantically coherent positioning of categorized elements enhance memory for the elements' names and locations (Saariluoma and Sajaniemi 1989, 1991, 1994). In general, the results give the experimental confirmation (regarding recall) to the common guidelines of interface layout design recommending spatial grouping of items from the same semantic category (e.g. Tullis 1997). The results also reveal some interesting interaction effects of spatial grouping and semantic coherence on the memory for the layout.

In the first experiment, semantic coherence was an explanatory factor but spatial grouping was practically irrelevant. The task was to learn the names of the elements. In the second experiment, spatial grouping was an effective factor. The explanation is the spatial task demands. It seems that the task demands determined the optimal organization of the layout. Task demands, either verbal or spatial, have also previously been shown to affect memory organization of words (Curiel and Radvansky 1998).

In the third experiment, the two attributes of a layout were both to be recalled. This makes the task significantly more difficult compared to the previous tasks. Evidently, learning the two dimensions concurrently create interference with each other. Recall of labels was more difficult than in the first experiment and recall of locations was harder than in the second experiment. Learning times show a similar pattern. This indicates that the association of label and location information in the same chunk requires specific processing, which is in harmony with the results of attention and memory research (Treisman and Gelade 1980, Styles and Allport 1986, Treisman and Gormican 1988).

Postma and De Haan (1996) have also found that the task of assigning items to locations is more difficult than location reconstruction. Assignment requires verbal resources, unlike location recall. This explains why the coherence effects in the third experiment do not immediately follow the outcome of the first and second experiment. The best learning results can be achieved when both spatial grouping and semantic coherence can be used. However, semantic coherence is clearly more advantageous than grouping. Spatial grouping in a semantically incoherent layout deteriorated recall of labels and locations. Presumably, the spatial grouping of labels creates false expectations about the simultaneous semantic coherence among the items. This indicates that if there are clear categories in which interface elements belong to, mixing them up in spatial groups may harm the user's performance more than random, spatially spread positioning.

The inferiority of spatially grouped, semantically incoherent layout in the third experiment contradicts the second experiment, in which the same layout was beneficial for location recall. Grouping facilitated the task significantly only when the layout was incoherent. Possibly the semantic categorization interferes with encoding or recalling locations when the layout is grouped semantically coherently. The user might process locations of clusters instead of single elements as in Hirtle and Mascolo (1986) study. Thus, whilst the user is more likely to encode locations of items when they are semantically related (McNamara and LeSueur 1989), the accuracy of location coding may suffer.

However, this does not explain why in Experiment 3, when both labels and locations had to be recalled, spatial-semantic clustering was beneficial also for location recall. Perhaps verbal factors outweigh the spatial ones in chunking, and thus layout organization facilitating verbal recall is more important when both label and location recall is required. Only when location recall is the main task, semantically incoherent and grouped layout would be more beneficial. This result may have interesting consequences for instance in complex icon search, in which users tend to rely on location recall (Blankenberger and Hahn 1991, Moyes 1994, Ehret 2002). In addition, mere grouping does not necessarily make the display less complex, but grouping should be 'good'. Spatial memory is improved if the locations form a good pattern in a Gestalt sense (Saariluoma and Sajaniemi 1989, 1991, Saariluoma 1992).

The main morale of these experiments is the importance of visual information chunking in interaction design. People use various dimensions of screen elements to group them and to construct visual chunks. This is understandable as it decreases memory lead, but in some cases, it may also have negative effects and impair the level of performance. When tacit rules created by chunks case unjustified expectations, the consequences are negative.

Chunking is one of the main mechanisms in human information processing. It is the very means people circumvent the limits of their working memory. Nevertheless, very little attention has been paid to visual information chunking and its role in searching information when interacting with computers. For this kind of reasons it is important to investigate and take into account the systems of tacit chunks users rely on in interacting with computers.

References

Attneave, F. 1955, Symmetry, information, and memory for patterns. *American Journal of Psychology*, **68,** 209–222.

Ben-Av, M. B. and Sagi, D. 1995, Perceptual grouping by similarity and proximity: Experimental results can be predicted by intensity autocorrelations. *Vision Research*, **35,** 853–866.

Bentum, K. 1998, Conceptual and spatial representations in memory. *International Journal of Cognitive Technology*, **3,** 39–44.

Blankenberger, S. and Hahn, K. 1991, Effects of icon design on human-computer interaction. *International Journal of Man-Machine Studies*, **35,** 363–377.

Byrne, M. D. 1993, Using icons to find documents: Simplicity is critical. InterCHI '93 (Amsterdam, The Netherlands: ACM).

Card, S. K. 1981, User perceptual mechanisms in the search of computer command menus. Proceedings of the first major conference on Human factors in computers systems (Gaithersburg, Maryland, USA: ACM).

Curiel, J. M. and Radvansky, G. A. 1998, Mental organization of maps. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24,** 202–214.

Decker, W. H. and Wheatley, P. C. 1982, Spatial grouping, imagery, and free recall. *Perceptual & Motor Skills*, **55,** 45–46.

Egan, D. E. and Schwartz, B. J. 1979, Chunking in recall of symbolic drawings. *Memory & Cognition*, **7,** 149–158.

Ehret, B. D. 2002, Learning where to look: Location learning in graphical user interfaces. SIGCHI '02 (Minneapolis, Minnesota: ACM).

Gittins, D. 1986, Icon-based human-computer interaction. *International Journal of Man-Machine Studies. Special Issue: Constructing user interfaces*, **24,** 519–543.

Glanzer, M. 1969, Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning & Verbal Behavior*, **8,** 105–111.

Green, A. J. K. and Barnard, P. J. 1990, Iconic interfacing: The role of icon distinctiveness and fixed or variable screen locations. INTERACT 1990 (Cambridge, UK).

Halgren, S. L. and Cooke, N. J. 1993, Towards ecological validity in menu research. *International Journal of Man-Machine Studies*, **39,** 51–70.

Han, S., Humphreys, G. W. and Chen, L. 1999, Uniform connectedness and classical gestalt principles of perceptual grouping. *Perception & Psychophysics*, **61**, 661–674.

Hirtle, S. C. and Mascolo, M. F. 1986, Effect of semantic clustering on the memory of spatial locations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 182–189.

Howard, M. W. and Kahana, M. J. 2002, When does semantic similarity help episodic retrieval? *Journal of Memory and Language*, **46**, 85–98.

Kemps, E. 1999, Effects of complexity on visuo-spatial working memory. *European Journal of Cognitive Psychology*, **11**, 335–356.

Magnussen, S., Dyrnes, S. and Nordby, K. 1997, Micro-position effects in visual short-term memory. *Scandinavian Journal of Psychology*, **38**, 139–142.

Marcus, A. 1995, Principles of effective visual communication for graphical user interface design. In R. M. Baecker, W. Buxton, and J. Grudin (eds) *Readings in human-computer interaction: Toward the year 2000* (San Francisco, CA: Morgan Kaufman), pp. 425–441.

McDonald, J. E., Stone, J. D. and Liebelt, L. S. 1983, Searching for items in menus: The effects of organization and type of target. Proceedings of the 27th Annual Meeting of Human Factors Society, October 10–14 (Norfolk: Human Factors Society).

McNamara, T. P. and LeSueur, L. L. 1989, Mental representations of spatial and nonspatial relations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **41**, 215–233.

Moyes, J. 1994, When users do and don't rely on icon shape. CHI '94 (Boston, Massachusetts: ACM).

Neely, J. H. and Balota, D. A. 1981, Test-expectancy and semantic-organization effects in recall and recognition. *Memory & Cognition*, **9**, 283–300.

Niemelä, M. and Saarinen, J. 2000, Visual search for grouped versus ungrouped icons in a computer interface. *Human Factors*, **42**, 630–635.

Palmer, S. E. 1992, Common region: A new principle of perceptual grouping. *Cognitive Psychology*, **24**, 436–447.

Postma, A. and De Haan, E. H. F. 1996, What was where? Memory for object locations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology. Special Issue: Working Memory*, **49A**, 178–199.

Saariluoma, P. 1992, Do visual images have gestalt properties? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **45A**, 399–420.

Saariluoma, P. 1994, Location coding in chess. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **47A**, 607–630.

Saariluoma, P. and Sajaniemi, J. 1989, Visual information chunking in spreadsheet calculation. *International Journal of Man-Machine Studies*, **30**, 475–488.

Saariluoma, P. and Sajaniemi, J. 1991, Extracting implicit tree structures in spreadsheet calculation. *Ergonomics. Special Issue: Cognitive ergonomics II*, **34**, 1027–1046.

Shneiderman, B. 1997, *Designing the user interface, third edition. Strategies for effective human-computer interaction* (Reading, Massachusetts: Addison Wesley Longman).

Styles, E. A. and Allport, D. A. 1986, Perceptual integration of identity, location and colour. *Psychological Research*, **48**, 189–200.

Teitelbaum, R. C. and Granda, R. E. 1983, The effects of positional constancy on searching menus for information. Conference on Human Factors and Computing Systems (Boston, Massachusetts, USA).

Toglia, M. P., Hinman, P. J., Dayton, B. S. and Catalano, J. F. 1997, The blocked-random effect in pictures and words. *Perceptual & Motor Skills*, **84**, 976–978.

Treisman, A. and Gormican, S. 1988, Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, **95**, 15–48.

Treisman, A. M. and Gelade, G. 1980, A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97–136.

Tullis, T. S. 1997, Screen design. In M. Helander, T.K. Landauer and P. Prabhu (eds) *Handbook of Human-Computer Interaction* (Amsterdam: Elsevier Science), pp. 503–531.

Wertheimer, M. 1923, Untersuchen zur lehre von der gestalt, ii. *Psychologishe Forshung*, **4**, 301–305.

Wiedenbeck, S. 1999, The use of icons and labels in an end user application program: An empirical study of learning and retention. *Behaviour & Information Technology*, **18**, 68–82.

Appendix

Table 8. The semantic categories used in the experiments

Predators	Flowers	Cars	Metals	Diseases	Drinks
Rodents	Berries	Numbers	Gemstones	Family	Liquors
Furniture	Spices	Colours	Cities	Occupations	Sports
Fish	Fruits	Mountains	Currencies	Instruments	Fabrics
Birds	Vegetables	Nations	Dances	Sciences	Foods
Root crops	Water birds	Garments	Planets	Tools	Games
Grains	Mushroom	Gases	Delicacies	Weapons	Music
Wood species	Types of waters	Month names	Head parts	Domestic animals	Internal organs