The AMODEUS Project
ESPRIT Basic Research Action 7040

Structuring Displays: a psychological guide

Jon May, Sophie Scott and Phil Barnard

12/5/95

UM/WP31

AMODEUS Partners:
MRC Applied Psychology Unit, Cambridge, UK (APU)
Depts of Computer Science & Psychology, University of York, UK. (YORK)
Laboratoire de Genie Informatique, University of Grenoble, France. (LGI)
Department of Psychology, University of Copenhagen, Denmark. (CUP)
Dept. of Computer & Information Science Linköping University, S. (IDA)
Dept. of Mathematics, University of the Aegean Greece (UoA)
Centre for Cognitive Informatics, Roskilde, Denmark (CCI)
Rank Xerox EuroPARC, Cambridge, UK. (RXEP)
CNR CNUCE, Pisa Italy (CNR,CNUCE)

Contents

About this Guide

This Guide is intended to help people who design computer displays to use psychological principles to choose the visual appearance of computer interface objects, their arrangement on the display, and their dynamic behaviour.

There are many books that provide 'guidelines' for designing displays – some tell you how and when to use different colours and typefaces, how to format columns and tables, and how to make your designs aesthetically appealing. This is not one of those books.

Although they provide a valuable service, and sometimes also try to explain why they are providing the advice that they do, guidelines are intended to be prescriptive – telling you what you should do for each part of a display. You can follow all of the advice that they provide for every individual part of your display layouts, and still find that you produce a design that is not 'easy to use'. Books of guidelines cannot tell you how to decide for yourself whether a display will be usable, nor how to identify the problematic parts of the design so that you can improve them. That is what this Guide tries to do.

It will introduce you to some psychological ideas about perception – the process by which people see objects in the world, recognise them and search between them. You don't need to be a psychologist to read this Guide – we've tried to avoid using psychological jargon – but when you have read it, you should be able to use these psychological ideas to analyse your display designs. The techniques this Guide teaches you will let you decide how difficult it will be for people to group objects together, to tell objects apart, to search for objects, and to switch their attention from one part of the display to another.

The Guide is organised into several sections. Each section introduces you to some ideas about perception, with some examples, and shows you how these ideas can be seen to affect the usability of display designs. The key points in each section are highlighted like this:

this is a key point, and a **special term** is shown like this.

These key points are summarised at the end of the Guide, so you can use these as an index to refer to particular issues. At several points in the Guide there are exercises for you to try, to check that you understand how the ideas can be used in practice.

The sections build on each other, introducing the simpler ideas first and the more complicated ideas later, and so this isn't a book that you can 'dip into', like a collection of guidelines might be. You have to read it through section by section – but when you have done that, we hope that you'll have learnt enough to put your new skills into practice, without needing to keep the Guide by your side.

Perception is active

Our visual environment is made up of objects rather than features. When we look around a room we see different objects, for instance, some books on a desk. It is hard for us to 'see' the pile of books as an area of different hues and shades, although that is what is represented by the pattern of light that is arriving at our eyes. The process of perception is one of structuring the sensory information that we receive from objects in the environment so that we can interact with them. We need to be able to see a set of differently coloured planes and surfaces as belonging to a single object, a book, that is distinct from the other planes and surfaces that represent the desk and the other books. If we pick up a book, we expect all of the parts of the sensory world that 'belong' to it to move together in a predictable way, and for all of the parts that belong to the 'desk' to stay where they are. If we try to pick up a stack of books, we know that the individual books might not remain as a stack, and that the stack cannot be treated in the same way as the individual books that it is made of.

These details about the structure of objects and their inter-relationships is not explicitly contained in the visual information. It must be interpreted, by combining the visual information with knowledge about the world, which we have learnt through our lifelong experience of interacting with it. This is why we can say that perception is active process, blending knowledge and sensation. The structure of the perceived world affects our interactions with it, and our interactions with the world affect our subsequent interpretations of its structure.

We use visual information to **perceive** objects within a scene

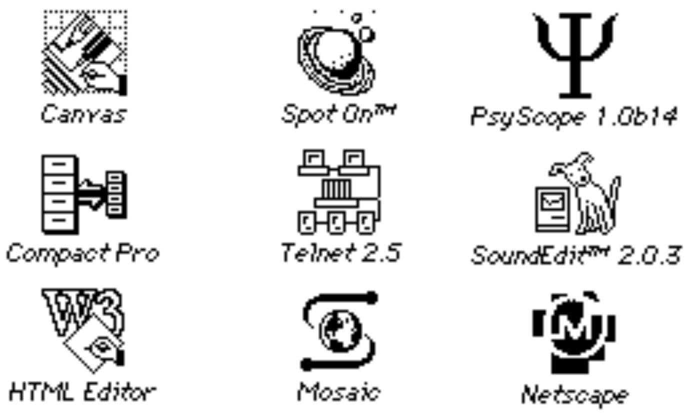


Figure 1: an array of different icons

Computer displays are just like the rest of the world in this respect, even though they are essentially two dimensional. Figures 1 and 2 show two groups of icons - one of these is a group of different icons, the other a collection of very similar icons. In Figure 1, if the user knows what a particular icon in the first array looks like, the dissimilarity will make it easier to locate - it may even seem to 'pop-out' from the array. But if they do not know what an icon looks like, and have to read all of the names, they may find it difficult to 'separate' the text labels from the 'background' of icons.

In Figure 2 the icons are all very similar, and so even if the user knows what the icon of the document they are searching for looks like, they may find it harder to locate than in the previous figure. However, the very similarity of the individual icons makes it much easier for the user to 'group' them as a single, ordered array, and for them to form a 'background' against which the text labels stand out. In this figure, the information provided by the appearance of each icon is less important than the fact that they cluster together to say "we're all documents": this becomes an attribute of the group rather than just an attribute of each separate icon.

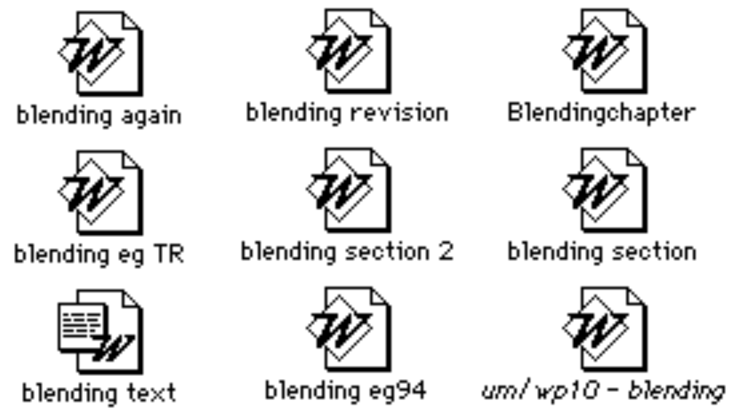


Figure 2: an array of structurally similar icons

There is nothing explicitly represented on the display in Figure 2 to tell the user that the icons form a group with a common attribute: no lines, boxes, colouring or shadows are needed. The similarity of the shapes, together with the knowledge the user has from their interface experience, is enough to implicitly inform them about the relationship. If this array were embedded within a real display, surrounded by other information, it would be easier for a user to find and attend to than the one shown in Figure 1.

Designing a computer interface is all about choosing the form of objects and arranging them within a two-dimensional area of the screen. As Figures 1 and 2 show, a correct choice of form and arrangement can affect the way that objects are perceived and dealt with by the user of the computer.

In this Guide we will explain some ideas that can help you to think about the way that people perceive visual scenes, and teach you techniques which you can use to analyse display designs.

The structure of visual scenes

Although computer displays are produced on two dimensional, flat screens, we use the same perceptual processes to perceive them as we do to perceive the real, three dimensional world. When we look at a visual scene, whether it is two or three dimensional, the features, colours, and textures in the sensory information that we receive from our eyes group together to form objects. The scene as a whole is a structure of objects. The objects have certain qualities - they stand out from their background and are discrete entities, which can often be named. If we look closely at an object, though, we can see that it also has a structure, and may be composed of other objects. We can perceive the world at several different scales, from a global level, down through many levels of detail.

Objects **group** together into larger objects, and can be **decomposed** into smaller objects

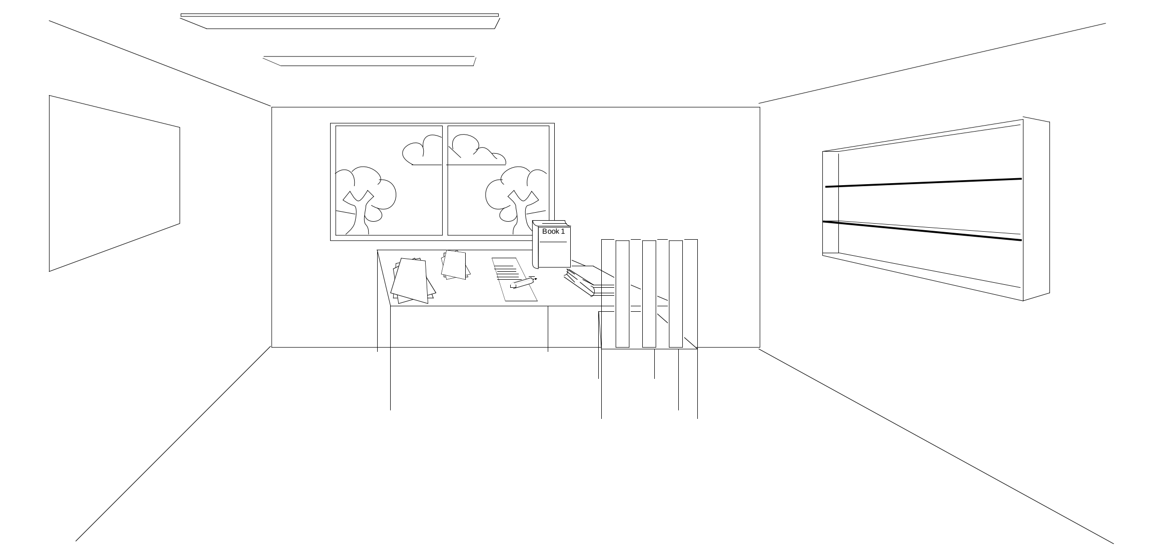


Figure 3: a visual scene – an office

You could stand at the door of the room in Figure 3 and see ‘an office - a room with objects in it’. You could then focus your attention towards the far wall, which is a plane surface with items of furniture superimposed on it. Within this level you could see the window, a chair and desk. Within the region of the desk you could see a pad of paper. The pad has a pencil resting on it, and is written upon. You could look at the text on the page by moving down into the structure of the pad, and moving down again you could see the individual words that make up the text (if you were near enough).

This hierarchy can be represented as a structure diagram, as in Figure 4, where the different horizontal groups in the figure represent different levels of visual structure. At each of these levels, sensory patterns of light are interpreted as forming a group of individual objects. Each object itself ‘contains’ visual details that can be further interpreted as another group of objects. The dotted lines indicate that some objects have further structural details that we have not included.

What we actually perceive from moment to moment is limited by the level at which we are analysing the scene. While attending to the pad of paper we can be aware of the relationships it has to the other objects within its own ‘group’ – the stacks of paper, and the books – and we can be aware of their shared relationship to the desk. We can also be aware that the pad itself has some structural details and, if we wanted to, could attend to some object of this structure; perhaps looking at a line of text.

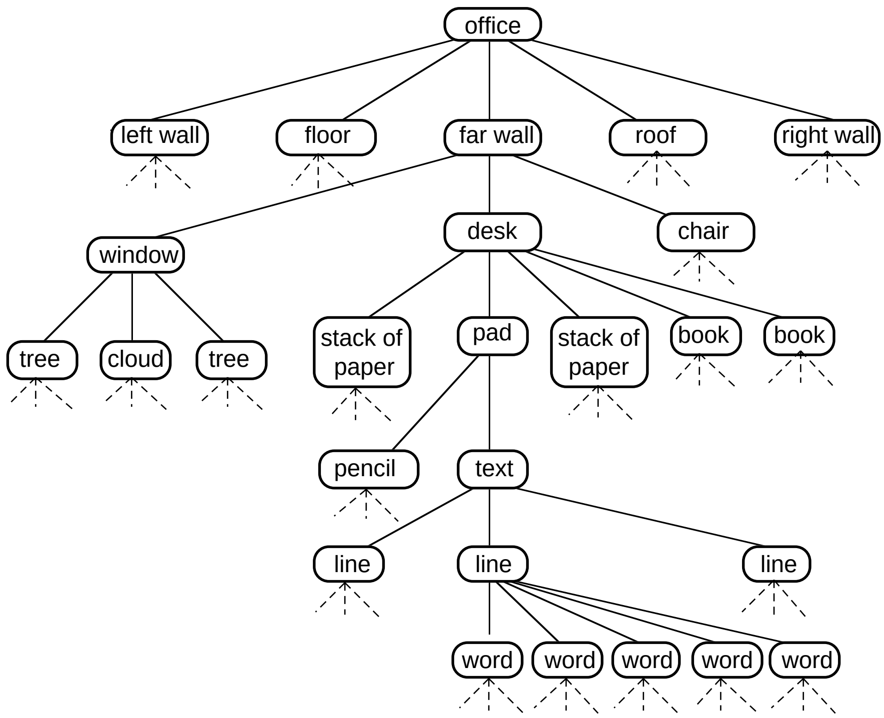


Figure 4: a structure diagram of the office scene

The hierarchical structure of the visual scene, as represented in Figure 4, constrains the direction of visual search. After having attended to the far wall, the words of text cannot be reached by looking at the structure of the window. We have to successively focus in to the desk, the pad and then the text before we can attend to the words. Likewise, after attending to a line of text, attending to a book requires a movement back up the structure, to an object that is at the same level of visual detail as the book (here, the pad of paper).

The **structure** of the scene constrains the way people can search through it.

These two ideas – the structure of visual scenes, and the transitions of attention between objects – are the tools that we will use to analyse the composition of displays. In general, a well composed display will be constructed so that the user can attend to the appropriate object easily. These tools help us to assess the ease with which a user can move their focus of attention around between objects. In the next section we will describe how they are derived.

Psychological subjects and transitions

To describe the way that we change the focus of attention, it is useful to think of the object that is being attended to as the 'psychological subject'. In the office example of Figure 3, there are several different objects on the desk. We can focus our attention on any of these objects, and we can shift our attention between them. Any one of them can be the psychological subject at different times. Other objects at the same level of decomposition in the visual scene form its context, and can be used to discriminate it from other similar or identical objects. Because these other objects provide information about the subject, they are collectively called its 'predicate'.

The object that is being attended to is the **psychological subject.** Other objects in the same group form its **predicate**.

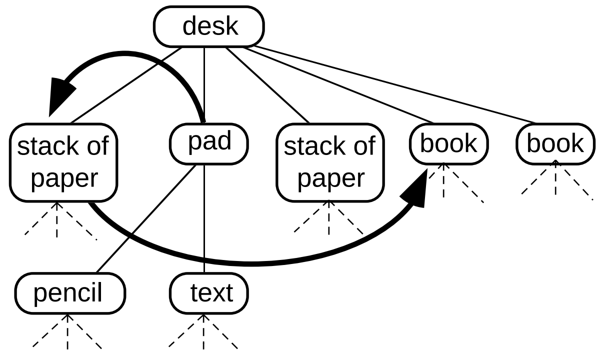


Figure 5: transitions in attention between objects on the desk

Figure 5 shows part of the office – the group of objects that are on the desk - as attention switches from the pad to a stack of paper, and then to a book, as indicated by arrows. Adding a lot of arrows to the structure diagram would make it rather complicated, especially if attention repeatedly moved back to the same object, and so we need to use a representation that can include time as a dimension.

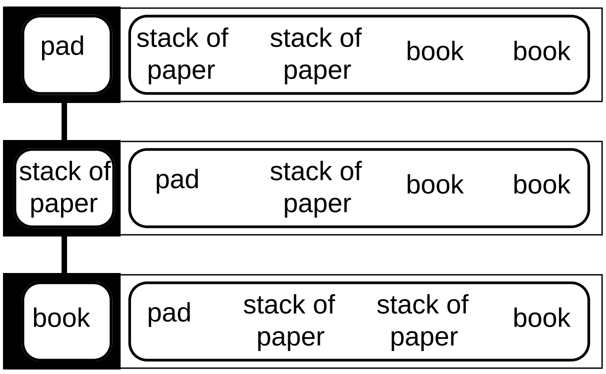


Figure 6: a transition path diagram showing the shifts in attention made in Figure 5

Figure 6 is an example of a 'transition path diagram' that describes the transitions in attention made in Figure 5. Each row represents a different moment in time, and a new focus of attention. One object is shown on a black background: this is the psychological subject at that moment. In the first row it is the pad, and the other objects form its predicate, and are listed in a group to its right. As successive transitions are made from object to object, each in turn moves left to become the subject, as shown by the second and third rows. The lines between the rows show the visual transitions that are made as attention shifts between the objects.

Structure diagrams show the hierarchical relationships between objects.

Transition Path diagrams show changes in the psychological subject and predicate in time.

Figure 6 might not seem to offer many advantages over Figure 5, but that is because the transitions were quite simple. As well as shifting attention between objects within a group, it is also possible to 'zoom in' and 'zoom out', attending to an object's group - the larger object it belongs to - or to a part of its structure - an object that it visually contains, surrounds, or is made up of. We need to be able to represent these possible transitions as well.

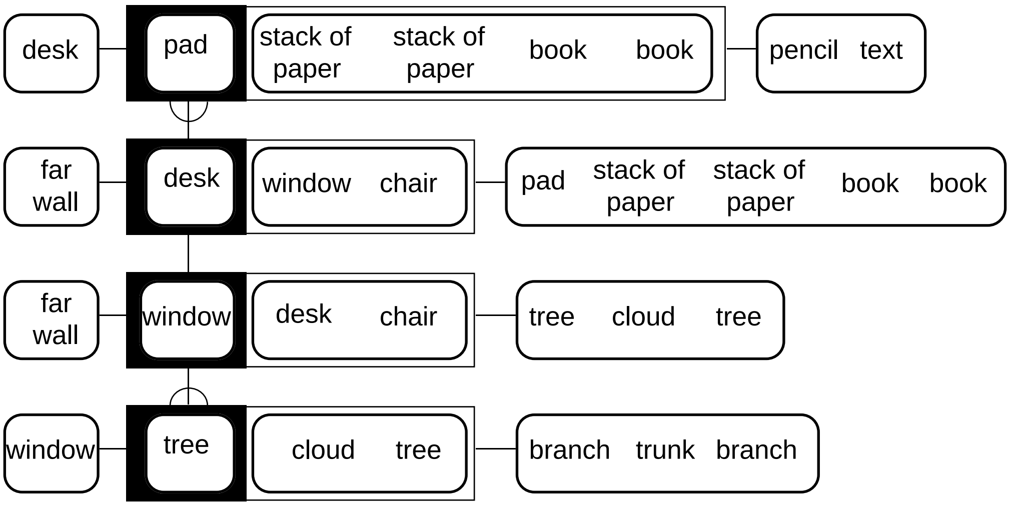


Figure 7: transitions as attention 'zooms out' and 'zooms in'

Figure 7 shows how these 'up' and 'down' transitions can also be represented in a transition path diagram. As well as showing the predicate of the psychological subject, each row includes (on the left) the group that the objects belong to, and (on the right) the constituent structure of the psychological subject. This diagram now contains all of the objects that could become the subject following a transition in attention. In the first row the pad of paper is again the subject, but the transition that is made next is 'up' the structure, to the desk. The 'fork' shape linking the first row to the desk indicates that the transition is 'up' the structure from many objects (the pad and its predicate) to a single object (the desk).

In the second row the desk is shown as the subject. Now the predicate consists of the other objects that are at the same level of decomposition as the desk - the window and the chair – and the ‘far wall’ is shown as the group that they belong to. The pad, the stacks of paper and the books, that were the active level of the previous representation, are now shown as the desk’s constituent structure. They have moved to the right, as has the desk. Each time a transition is made ‘up’ a structure, the old group moves right to become the new subject, and the old subject-predicate level moves right to become the new constituent structure.

In the third row a ‘within level’ transition is made to the window, so that it becomes the subject: the group remains the same, but the constituent structure changes, to show which transitions ‘down’ the structure are now possible from the window. The previous subject has become part of the predicate. A point to notice here is that the objects within the predicate are ‘unordered’ - they are all equally able to become the subject. The second and third rows are linked by a plain line, to show that the transition is just from one object to another, within the same group.

Finally, in the bottom row a transition is made to a tree – one of the objects that the window ‘contains’ in its structure. The group of objects that was on the right of the third row has moved left to become the ‘active level’ of the representation in the fourth row. The subject is the tree that is being attended to, the predicate consists of the other objects in the window’s structure, and the tree’s constituent structure must be included on the right of the row. The window has also moved left, to become the group. Each time a transition is made ‘down’ a structure, the old subject moves left to become the new group, and its constituent structure moves left to become the new subject-predicate level. The ‘fork’ linking the third and fourth rows now indicates that the transition has been from a single object, the window, to the many objects in its structure.

Transition path diagrams help to make it clear how simple or how complicated it will be for users to move their attention from object to object within a display. On each row, all of the objects that could be attended to following a transition are indicated. A transition ‘up’ the structure makes the group and subject-predicate move right in the row. A transition ‘down’ makes the subject and its constituent structure move left in the row.

In Figure 7, it took three transitions to look up from the pad, and to look at a tree. It might take a user of a computer several transitions to move their focus of attention from the document they are reading on-screen to locate an icon in a menubar, depending on the structure and grouping of all of the objects.

In analysing a display, it is helpful to construct a structure diagram first, and then to use it to draw transition path diagrams for particular tasks that a user will want to carry out. The next section shows how this can be done for a typical computer display.

Using diagrams to analyse a display

The ‘office’ example was a real-world, three dimensional structure, but the structural and transition path diagrams can be used to analyse two dimensional computer displays. The only differences between the two ‘control panels’ in Figure 8 are the boxes that have been drawn around the groups of words and buttons. This might be an aspect of the design that is left to a designer’s aesthetic judgement, or it might be constrained by the ‘toolbox’ used for building the display from software objects.

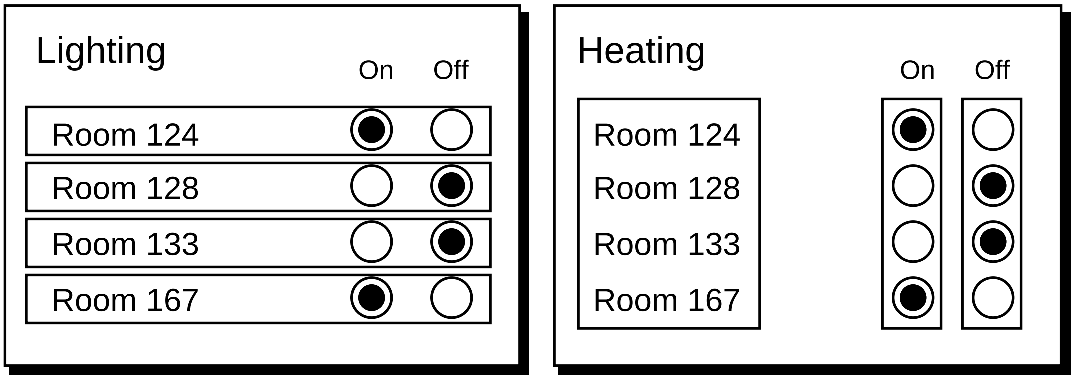


Figure 8: two ‘control panels’ that differ in the way the objects have been grouped by the designer

The ‘lighting panel’ has boxes that relate objects together functionally, so each room label is linked to its own on and off buttons. In the ‘heating panel’, the objects have been linked by type, so that all of the room labels, on buttons and off buttons each form different groups. This is a fairly small difference, and if anything, the heating panel looks more aesthetically appealing. The structure diagrams for the two panels (Figure 9) show the difference that these boxes make to the grouping of the objects. The lighting panel is made up of four ‘groups’, one for each room, each containing a room label and an empty and filled button. The heating panel consists of three groups, one of four similar room labels, and two groups of mixed circles.

If we draw transition path diagrams for a user who has to turn the lights and heat on in Room 133, we can see the difference that these groupings have made to the panels’ ‘ease of use’. For the lighting panel, the task is quite simple, as the diagram in Figure 10 shows. The button that turns the lights on is part of the predicate of the label for ‘room 133’ and so only one transition is necessary. The transition is made within a single group, and so the object to the left of the subject does not change. The empty circle has no constituent structure, and so when it is the subject nothing is shown to the right of the row, to indicate that no further transitions could be made ‘down’ the structure.

The situation is quite different for the ‘heating panel’. Now the room label and the button are in different groups, and the user has to momentarily move their attention up the structure to the ‘rooms’ group, across to the on-buttons, and then down again to the third button (Figure 11). Here three transitions are needed instead of one, and so for this particular task, we can say that the ‘lighting panel’ will be easier to use than the ‘heating panel’.

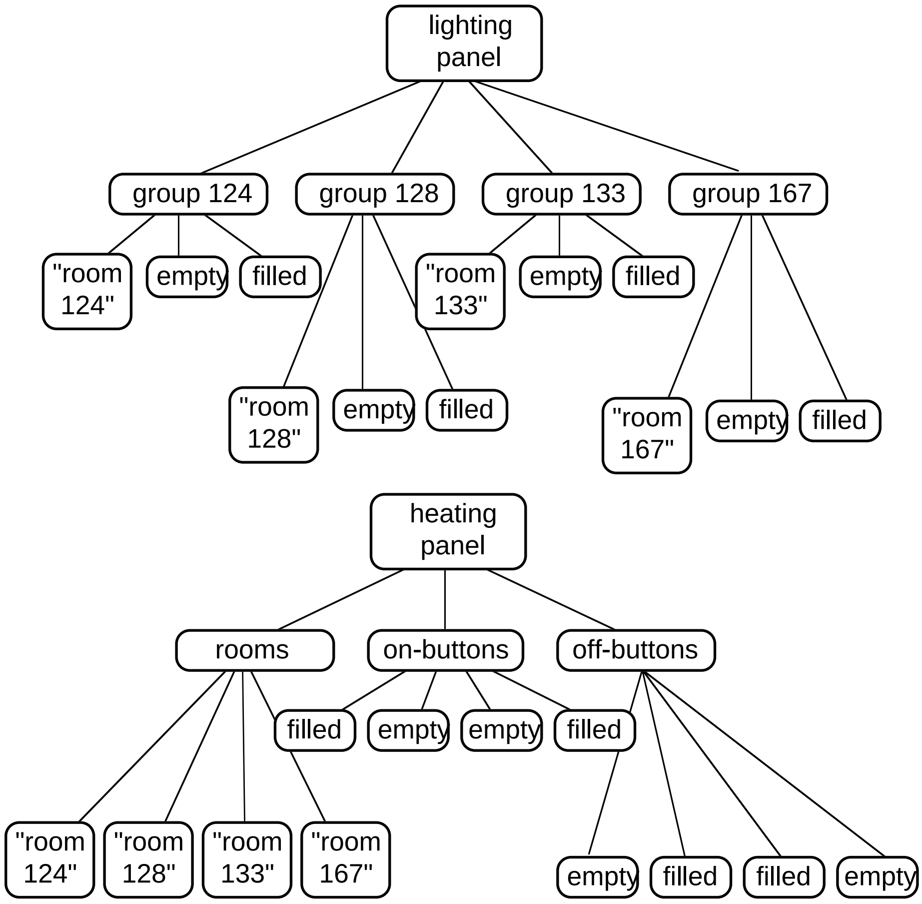


Figure 9: structure diagrams of the two control panels

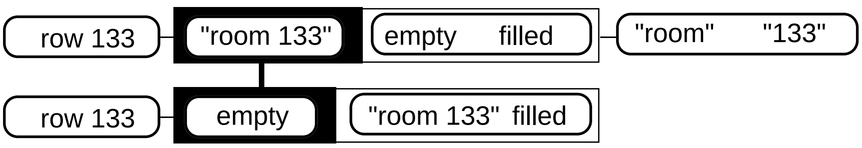


Figure 10: a transition path diagram for the lighting panel

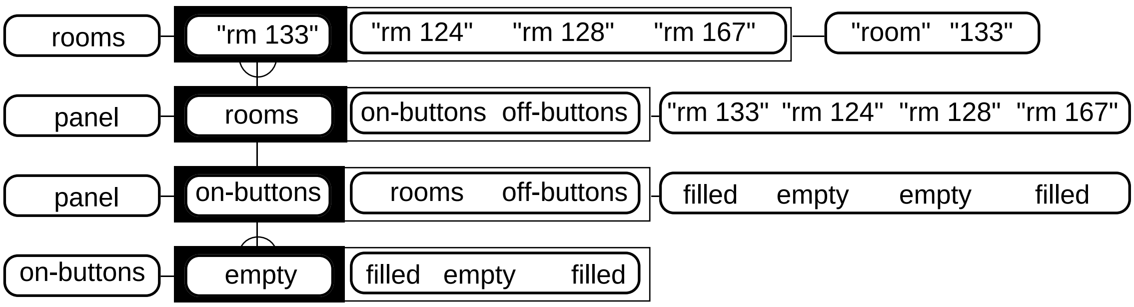


Figure 11: a transition path diagram for the heating panel

The idea of ‘task’ is very important, of course. If instead of a task that required the use of the room label and a button, the user had to operate on each of the on-buttons in sequence, regardless of the room labels, the ‘heating panel’ might be found to have an advantage. Suppose the user just had to make sure that all the heaters were on. Once the user had located the group of on-buttons, and attended to one of the buttons, the other buttons would all be part of the predicate. This task would require fewer transitions than the equivalent task of turning on all of the lights.

This example shows that it is vital to make the grouping of screen objects correspond to the task that the user is going to perform, because this determines the way that they will have to move their attention between the objects. In choosing between different possible forms for objects and different ways of arranging them, the designer is attempting not just to make an aesthetically pleasing interface, but one which helps the user perform a particular task.

In this example the grouping was done fairly obviously, with boxes, but there are a number of perceptual tricks that can be used to make objects form groups. In the rest of this Guide we will illustrate some of these methods, explain why they work, and show how the structure diagrams and transition path diagrams can help you ensure that the grouping of objects on the display matches the user’s task.

Objects and groups

In Figure 8, objects were grouped explicitly, with boxes. Objects can also be grouped according to their appearance or spatial arrangement. Figure 12 shows how four triangles can be grouped in different ways due to their:

- proximity (being very near to each other)
- sharing a colour
- sharing a boundary
- sharing a junction.
- collocation (being superimposed or intersecting)

These are 'physical' relationships that can be derived from the visual information, and in some cases the groups that result appear more 'obvious' than the original triangles.

Visual information can affect the way that objects form groups

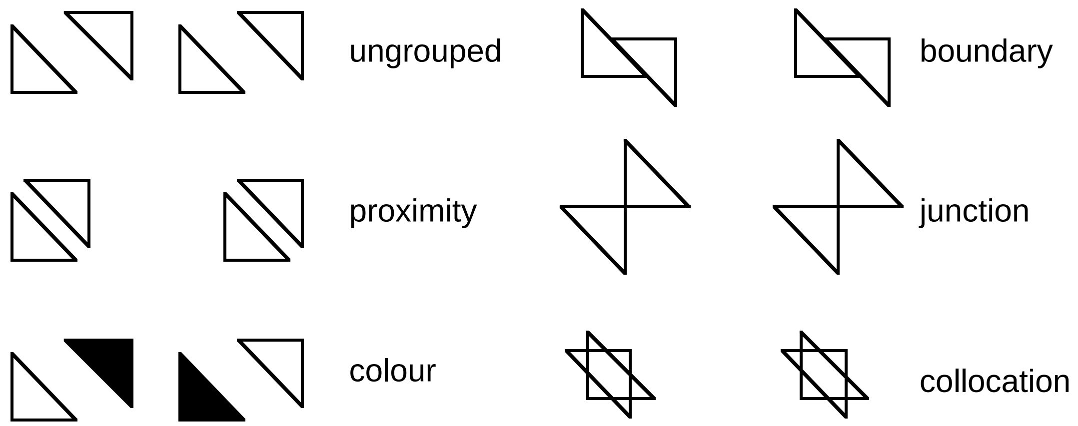


Figure 12. Primitive relationships between visual features affect grouping.

If you had to describe the 'junction' or 'collocation' parts of this figure, without having seen the rest of it, you would probably not make any mention of triangles. Instead of calling the junction group 'two triangles joined at one corner', you might call it 'an hour-glass' or 'a bow-tie'; and you might call the collocation group 'a six-pointed star, slanting backwards'. In both of these instances, you are describing the 'group' that is composed of the two triangles, and not the triangles themselves. We can take this point even further, because even in the 'ungrouped' instance, we perceive 'triangles', and not individual horizontal, vertical and oblique lines. In terms of the basic processes of visual perception, it is even arguable that we actually 'see' lines as 'end-points', 'corners' and 'middles' – but while this may be what we 'see', it is clearly not what we 'perceive'.

Figure 13 summarises the process of perception that we have described so far. Sensory information about the world is detected by the eyes, and turned into a 'visual representation' that contains a wealth of detail about colours, shades, contrast, angles and edges. A mental process then 'interprets' this information, transforming it into an 'object representation', which contains information about lines, shapes, depth, position and orientation. The object representation is then picked up by other mental processes, to be used as the basis of our interaction with the world. The structure diagrams and transition path diagrams portray the object representation, not the visual representation.

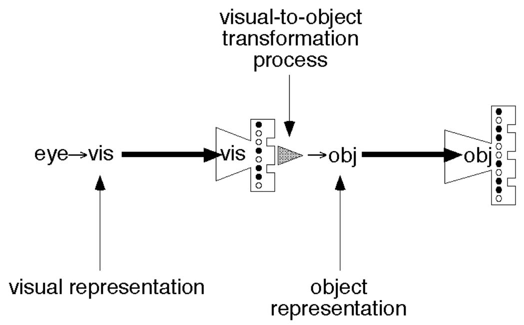


Figure 13: a visual representation produced by the eyes is transformed into an object representation

It is important to remember this distinction between the sensory level of information in the visual representation, and the perceptual level of information in the object representation. One advantage of making this distinction is that it helps us analyse what people will *subjectively* think about a display design (their object representation) as well as what is *objectively* presented to them on the computer screen (their visual representation).

Sensory information from the eyes forms a **visual representation.**

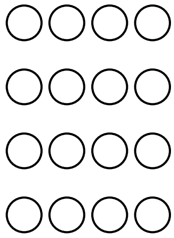
Perceptual information is contained in an **object representation.**

Structure diagrams and Transition Path diagrams portray the object representation.

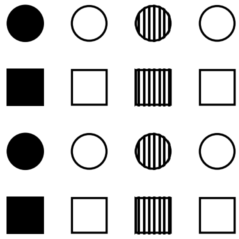
Exercise Set 1

1 Draw structure diagrams for these sets of objects – start each with a group called ‘set’ at the top level of the structure and use up to four levels.

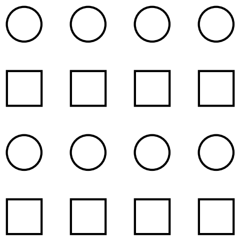
2 Describe the visual attribute that ‘causes’ each group.



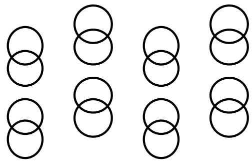
(1)



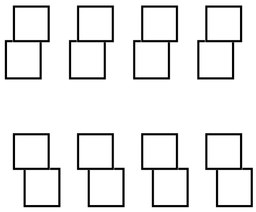
(2)



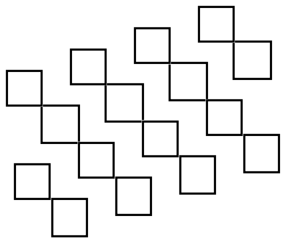
(3)



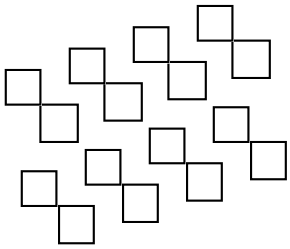
(4)



(5)



(6)



(7)

Psychological Subjects pop-out

One way to approach the problem of deciding when objects form groups and when they don't is to consider the phenomenon of 'pop-out'. This happens when there are several objects forming a group, and one object that doesn't join the group. In Figure 14, for example, there are three groups of circles. If you just look at the left hand group, all of the circles are exactly the same, and none of them stand out any more than the others. Because of their 'proximity', they are all members of the same group, and if we were to draw a structure diagram, each circle would be represented at the same 'level' of the structure as the others (shown in the left hand part of Figure 15).

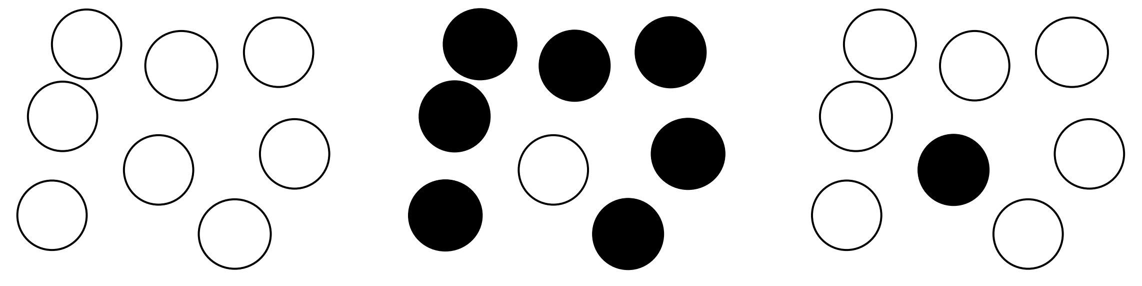


Figure 14: pop-out of the psychological subject

The central and right hand groups are different. In these groups all of the circles are the same size, but one of them is a different colour. You have no difficulty in noticing which one it is, because it seems to 'pop-out' from the others – and the structure diagrams for these two groups, shown in the central and right hand parts of Figure 15, represent this.

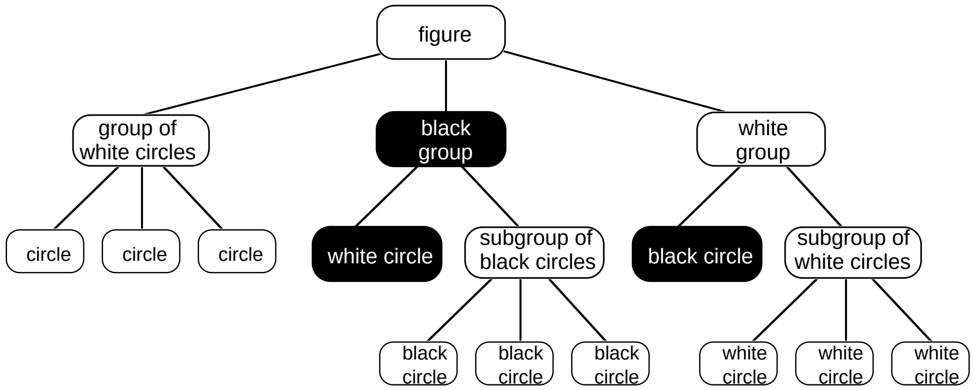


Figure 15: structure diagrams of figure 14

The circles that share the same colour all form a subgroup, to which the different circle does not belong. Both the subgroup of similar circles and the different circle are part of a larger group, and so we see them as related, but when we focus on the whole group to see the objects that it is made up of, we perceive the single circle and the group of circles.

Objects that are spatially close to a group, but not part of it, seem to ‘**pop-out’**

If we drew a transition path diagram for someone viewing the central group, we would show them attending to the figure as a whole, with the central group as their psychological subject. This would be followed by a transition to its structure that made the white circle the psychological subject, and the subgroup of black circles its predicate. In this case, the effect of pop-out is so strong that it almost forces the viewer to make the white circle the subject as soon as they attend to the structure of the central group. Even if the viewer wants to look at one of the black circles, they have to attend momentarily to the white circle, and then make additional transitions to the subgroup, and then into its structure to find a black circle.

It doesn’t matter whether the odd-one-out is black or white: as long as it is different, it becomes the psychological subject and pops-out, simply *because* it is different. As with the triangles in Figure 12, there is nothing in the sensory information that explicitly tells us that the black circles in the central part of Figure 14 all form a group, to which the white circle does not belong. At a visual level, they are all just areas of varying colour. The structural information that relates them together as members or non-members of groups is added by the process that transforms the visual representation into an object representation.

The **visual-to-object transformation** adds structure to the sensory information

If four of the circles in a group were black and four were white, then there would be two equally sized subgroups and neither would pop-out. If there were subgroups of five and three circles, the effect would not be as strong, but it is likely that the smaller subgroup would form the psychological subject, and the larger group would be its predicate. The process that transforms the visual representation into an object representation ‘favours’ the part of the visual scene that is different, and produces representations organised with them as the psychological subject.

In Figure 15 we have drawn the ‘different’ circle in each group in white text on a black background, to show that it pops-out. At the level above, we have indicated that the ‘black’ group pops-out in the same way, because it is different in ‘colour’ to the other two groups. When you look at the ‘whole figure’, the black group pops-out; but once you have attended to it, its white circle pops-out. Of course, it isn’t just colour that can make things different to their neighbours. Figure 16 shows that pop-out can also happen for shapes.

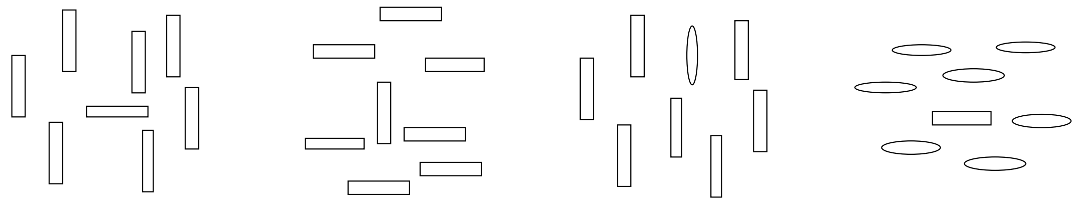


Figure 16: pop-out of shapes

Again, it doesn't matter why the shape is 'different'. You might say that the oblongs are all the same, and have just been rotated, but this is enough to make them different in the visual representation. As long as a shape is different, the transformation from a visual to an object representation picks it out as the psychological subject, and the other objects form a group that becomes its predicate. If you think about looking for objects in the real world, this bias of perception makes sense: more often than not we are searching for objects against a background, looking for one particular object that is different to the rest of the scene. Whether it is a ripe red apple in a tree of green leaves and green apples, or an icon on a computer screen, it often has some visual feature that makes it stand out from the background.

While higher mental processes could spend time and energy making transitions through a representation to locate the correct object, it is generally economical for the visual-to-object transformation to pick up the implicit information from the sensory data and to make the odd-one-out the one that gets attended to first. We make use of this tendency in the transition path diagrams, by drawing the psychological subject against a black background: it immediately pops-out from the diagram and orients you to the part of the figure we are likely to be describing.

We have also been using this convention in the structure diagrams, to indicate an object that pops-out to become the psychological subject. Unlike the transition path diagrams, which indicate the object that actually is the psychological subject at any moment, the structure diagrams aren't showing 'processing', but just the structure. The objects that pop-out aren't always psychological subjects, but will be if their level is attended to. To distinguish these 'potential' psychological subjects from 'actual' psychological subjects, we'll use the term 'pragmatic subject' – this means that the object can be expected to become the psychological subject for pragmatic reasons.

The **pragmatic subject** is the object that will become the psychological subject when the structure of its group is attended to, because of its visual features.

Objects that are **different** to others in their group can be pragmatic subjects.

Exercise Set 2

1. Draw transition path diagrams for Figure 14, showing the transitions that are needed to look at a white circle in each subgroup (you can base them on the structure diagrams from Figure 15).
2. Draw a structure diagram for Figure 16, indicating which object forms the pragmatic subject of each group.
3. Using this structure diagram, make transition path diagrams to show how a ‘horizontal oblong’ would be located in the first, second and fourth group (there isn’t a horizontal oblong in the third group!)

Pop-out of groups

Colour and orientation aren't the only sensory cues that the visual-to-object transformation can used to pick out part of a scene as a pragmatic subject. Other attributes can also be used – but in a different way. In the two arrays in Figure 17, the size of the circles is varied.

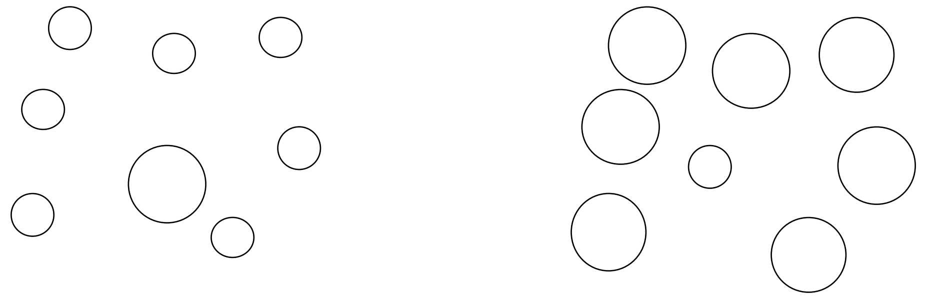


Figure 17: asymmetry of pop-out for different size

Although it is still easy to find the small circle among the big ones, it is not quite as easy as finding the big circle amongst the small ones. The big circle amongst the little circles is the pragmatic subject of its array, but the group of big circles is the pragmatic subject of the other array (Figure 18). To make the small circle the psychological subject, a transition from the group is needed. While colour and orientation were symmetrical (black and white being equally able to pop-out) and it was an object's 'difference' that was the cue, here the attribute is 'asymmetrical', and the visual-to-object transformation always favours the larger-sized objects.

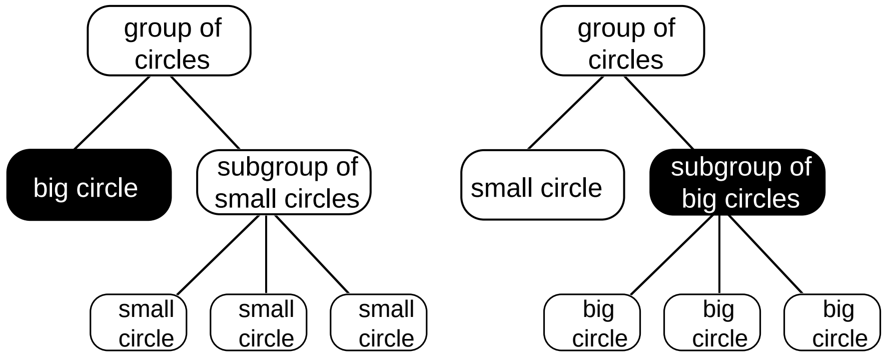


Figure 18: structure diagrams for figure 17

The same asymmetry can be seen with the length of lines in Figure 19. The reason for this asymmetry is that the size of a visual object is related to its closeness to us – in general, the larger an item is, the nearer it is. The visual-to-object transformation is now choosing the closer object as more likely to be of interest, and so makes it the pragmatic subject. Again, this seems to make sense in terms of the real world: if you are in a tree picking apples, the ones that are visually larger are more likely to be within reach than the ones that are visually smaller. The same rule of thumb applies to contrast and brightness, since as things get closer to us they reflect more light, and are less obscured by anything that is in the air.

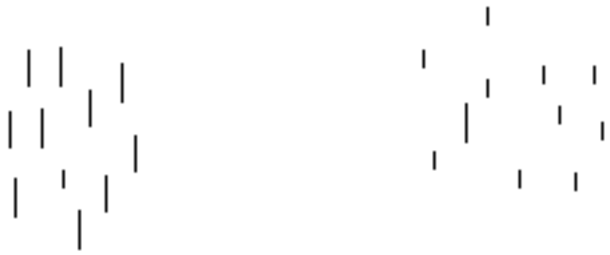


Figure 19: asymmetry of pop-out for line length

In many situations computer interface designs can take advantage of this bias towards difference and nearness. Like the white-on-black convention that we have adopted for representing the psychological and pragmatic subject, words and icons that are 'selected' usually become highlighted in some way, partly to provide feedback about the selection, but also to make sure that the user is actually attending to the part of the display that they have acted on.

Objects that are larger or brighter appear **nearer** and can be pragmatic subjects.

Options on menus and in dialogues that are unavailable are shown 'greyed out' by reducing their contrast – this indicates their unavailability, and also makes them less likely to be attended to, since they will no longer form part of the group that is the psychological subject when the menu or dialogue is viewed. In Figure 20, for example, a set of commands that operate on Tables in a word-processor are greyed out when the user has selected an ordinary paragraph of text – the other paragraph formatting commands are still black, and so they form a pragmatic subject that immediately grabs the user's attention.

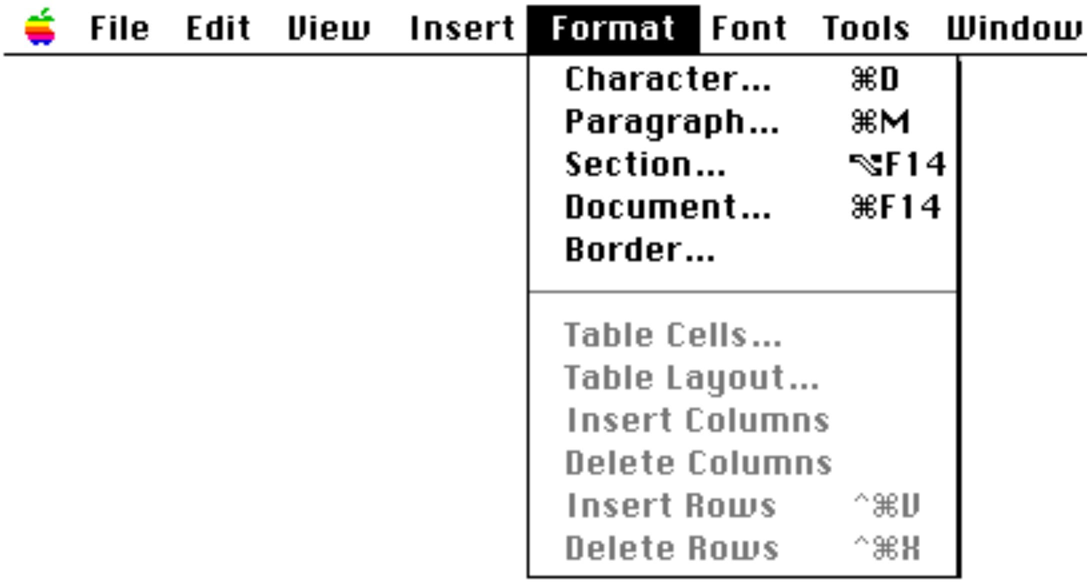


Figure 20: the 'greyed out' options are unavailable, leaving the black options as a more salient group

Use of these attributes can also help people to discriminate objects by guiding their attention to the part of its structure that distinguishes it from other objects. Figure 21 contains an array of four abstract icons, each of which is made of a diagonal cross and an upright cross. In (a) both crosses are shown by lightly dashed lines, but in the others one of the crosses is drawn as a bold, solid outline. Looking at each icon in turn it is clear that these changes affect the way that they are interpreted.

The icon (a) could be seen as two dashed crosses superimposed on one another at an angle, as four dashed lines, or as an eight pointed star. To its right, icon (b) has one dashed cross and one solid diagonal cross – the size of this cross makes it more salient, as it forms the pragmatic subject and the dashed upright cross becomes its predicate. The next icon, (c) has the same visual structure, but has been turned through 45°. This change in orientation is sufficient to render this solid, upright cross as an object that is different to the solid, diagonal cross of (b). Finally icon (d) has a large black diagonal cross – again the size of the diagonal cross makes it the pragmatic subject of this icon, but its colour also makes the icon that it belongs to likely to form the pragmatic subject of the whole array.

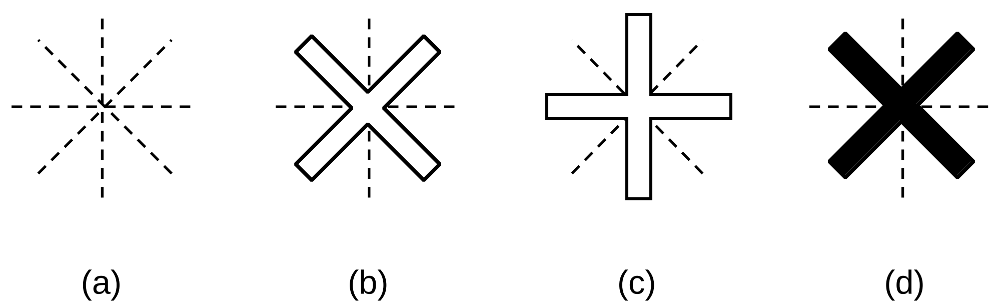


Figure 21: four abstract icons

If this array is attended to, it is probable that (d) may be the pragmatic subject, as its black solid cross is both 'nearer' than the thinner dashed crosses and 'different' to the solid, white crosses. The nature of the group of crosses that is icon (d) is defined predominantly by the nature of its pragmatic subject – because if it is looked at, the pragmatic subject is the first part of its structure that will be attended to.

The 'appearance' of an object is determined by its pragmatic subject

Exercise Set 3

1. Draw transition path diagrams for the location of a big circle and a small circle in each part of Figure 17.
2. Draw structure diagrams for both parts of Figure 19, and transition path diagrams for the location of a small and a large line in each part.
3. Draw structure diagrams for each of the icons in Figure 21.
4. Draw transition path diagrams for Figure 21, showing the transitions necessary to attend to the diagonal crosses of each icon. Which cross is hardest to attend to?

Pragmatic Subjects and Icon search

As the number of icons on an interface increases, and the range of functions that have to be represented proliferates, there is a tendency to design the icon to 'represent' the function in an almost pictorial way. This has a clear advantage when the icon is presented to users on its own, because it is easy for them to 'see' the relationship between the icon and its function (Figure 22). What is not so clear-cut is the effect upon the icons 'findability'.

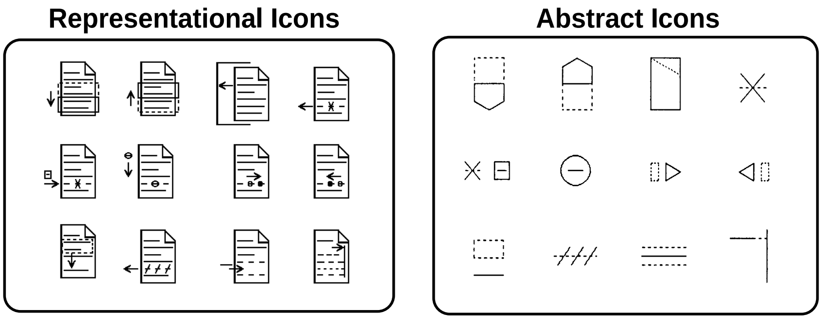


Figure 22: the representational and abstract sets of word-processing icons

Figure 22 shows representational and abstract icons that have both been used to stand for the same set of word-processing commands. The representational icons all look like pages of a document, with lines of text and arrows or boxes indicating the result of their function. The abstract icons are much simpler, and although they too provide some sort of semantic link between their appearance and their function, you really have to know what the possible functions are to work out what each icon might do.

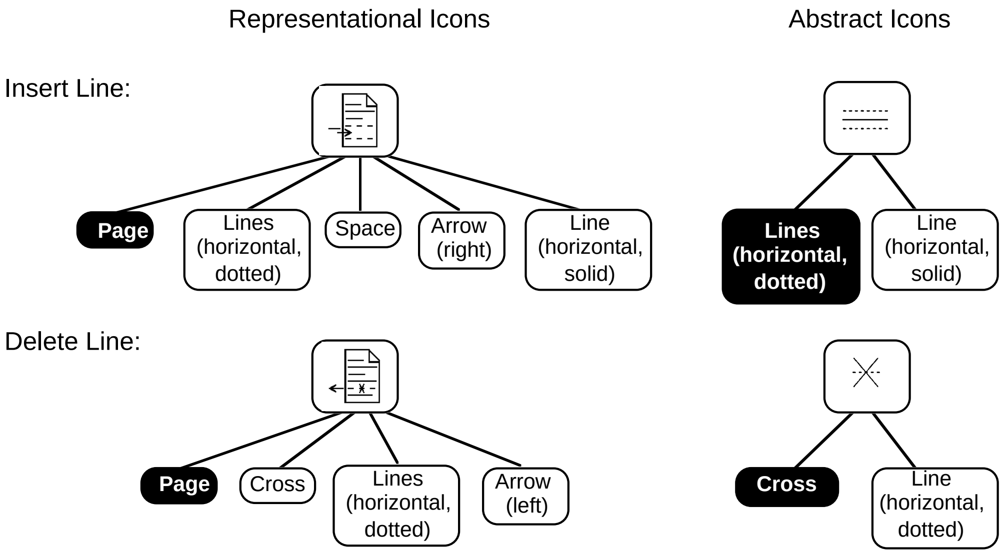


Figure 23: structure diagrams of two representational and abstract icons

In experiments where the position of the icons in the array varied, people using the representational set took longer to find the one they wanted than did people using the abstract set. If the icons were kept in the same position from trial to trial, so that users could remember the rough location within the array of each icon, and could 'look' straight for them without searching, the differences between the icon sets narrowed markedly.

The structure diagrams in Figure 23 show four of the icons from Figure 22: the two from each set that represented 'insert line' and 'delete line'. The icons from the representational set clearly have a more detailed structure than the icons from the abstract set, but they also have the same pragmatic subject as each other. To tell them apart, the user has to attend to their predicate as well as to the subject.

When the time that it took people to find each icon was compared with its internal structure, a clear relationship was found. The greater the degree of similarity that the icon's structure had to other icons in its set, the longer it took to find an icon. This suggests that users form a mental image of the 'target' icon that they are looking for, and then compare 'candidate' icons from the array with this internal mental image to see if they matched. An icon in an array would be a candidate if it had the right pragmatic subject, and it would match the target if it also had the right predicate. Icons that needed more objects of their predicate to be evaluated to be discriminated from other potential candidate icons would take longer, overall, to locate.

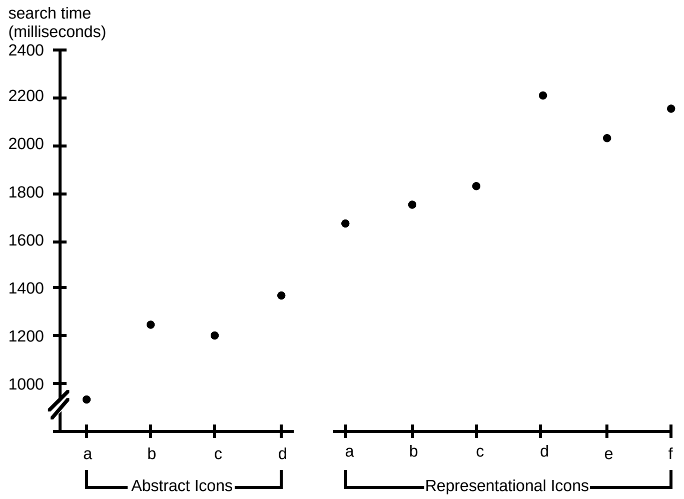


Figure 24: the more complex the discrimination, the longer it took to find an icon.

This helps explain why the representational icons took longer to find – it wasn't because they were representational, but because they were all so similar – and even within the sets, it was possible to show that the more complex the discrimination, and the more candidates that shared its pragmatic subject, the longer it took to locate an icon (Figure 24). In the abstract icons the pragmatic subjects are mainly different, which means that the search can be carried out at the level of the icon, using the more salient information.

People can search for an object by looking for objects that have its pragmatic subject

You might remember a similar effect of the pragmatic subject from the very first figures in this guide (Figures 1 and 2). Figure 25 again shows an array where all the icons in the window have different pragmatic subjects.

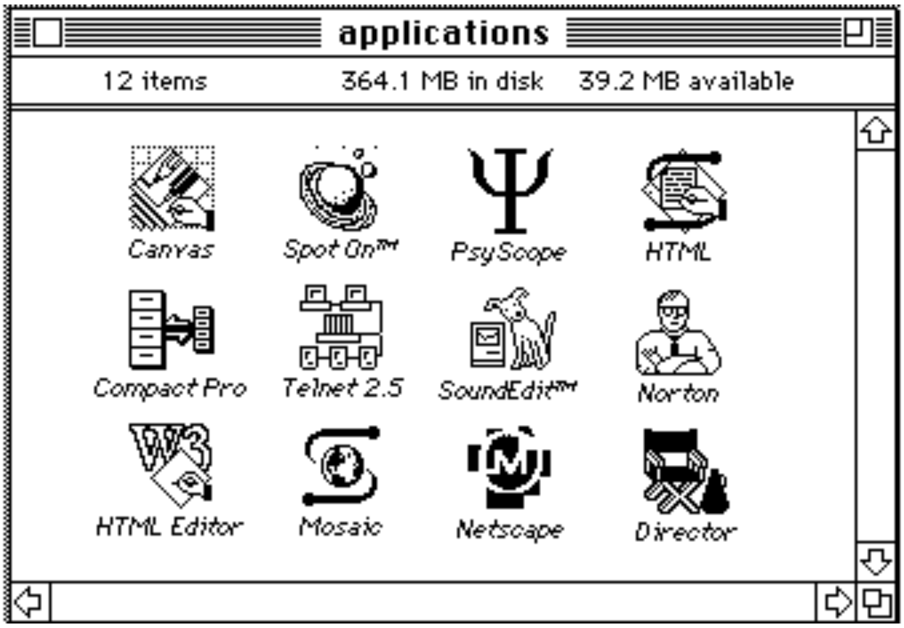


Figure 25: icons with different psychological subjects, in a Macintosh Window

These icons form a group of icons in an array, but their visual structures do not lead the user to see them as forming any subgroups. When an icon is searched for in this array, the icons can be discriminated from one another by their pragmatic subjects, without their structure needing to be evaluated. As you look from icon to icon in this array, you make the visual transitions represented in Figure 26.

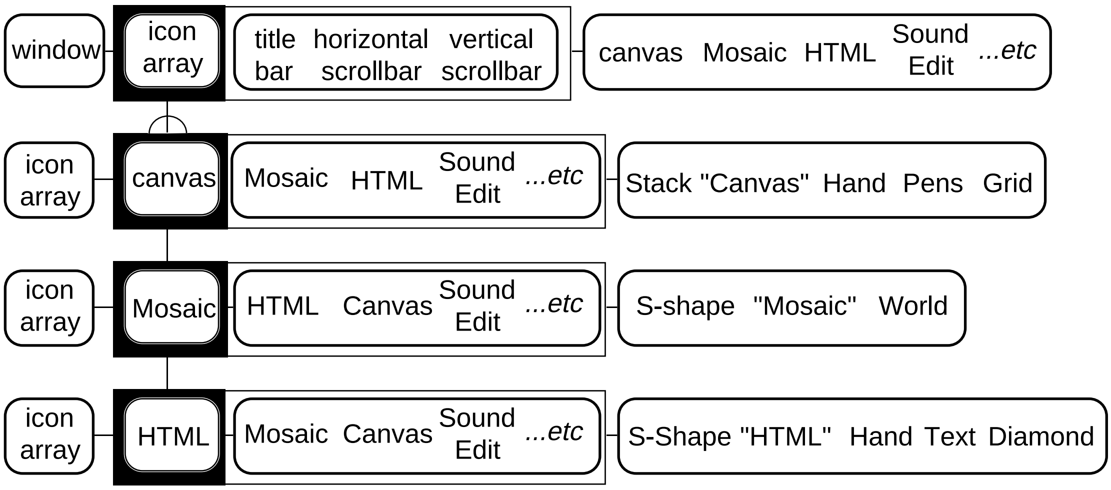


Figure 26: a transition path diagram showing the search for an icon in Figure 25

None of the icons in this array have any real advantage over each other: if you ‘know’ what icon you are searching for, and can form an ‘object’ image of the target, then you can probably locate it quite rapidly. Try finding the icon in Figure 25 that looks like a dog sitting next to a Macintosh computer. In contrast, try finding the icon in Figure 27 that contains a picture of a Macintosh computer. Now the subjects of many of the icons are similar, and you have to evaluate more information about each icon, as with the representational word-processing icons of Figure 22. The corresponding transitions are shown in Figure 28.

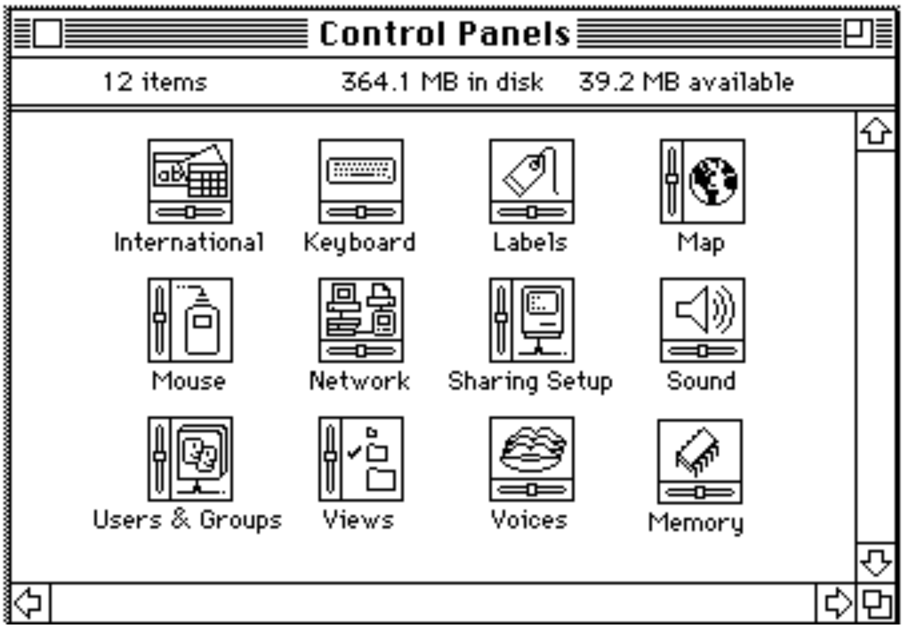


Figure 27: an array of different icons that share psychological subjects

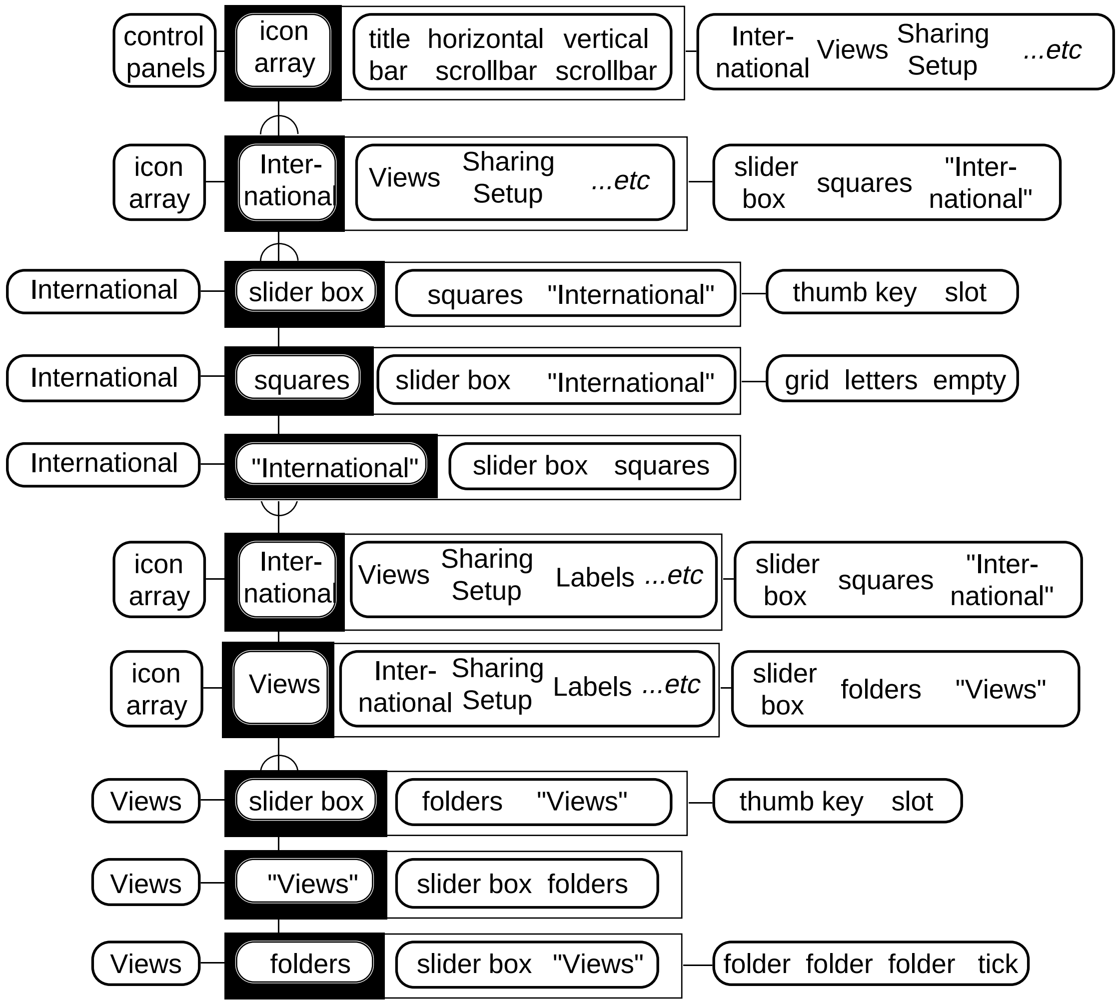


Figure 28: the transitions required to locate an icon in the ‘Control Panels’ array of Figure 27

The pragmatic subject of each icon is the ‘slider box’ that surrounds each icon’s contents, and so it is not possible to ignore them and locate the Macintosh directly. Each time an icon is attended to, a transition must be made away from the pragmatic subject to examine the rest of the icon’s contents. Again, none of the icons have much of an advantage over each other, but this array is harder to search than the one shown in Figure 25, because more transitions are required to search through its structure. In Figure 28, two incorrect icons are searched before the correct one is found – in an array of twelve icons, the average number of icons that would be evaluated in this way would be 6.5!

As with the circles of different colour, and the lines of different orientation, when one icon in an array has a completely different pragmatic subject to the others, pop-out happens, and that particular icon is very easy to find - this is shown in Figure 29. The icon that does not belong to the group of ‘Word’ document icons pops out from the array. Even though the Word icons are not all identical, and have different text labels, they are difficult to search through. The Word icon that is ‘different’ to the others still has the same subject (the document shape) and also shares a predicate object (the large ‘W’) and so is nothing like as easy to locate as the ‘ψ’ icon.

A general conclusion that you can draw from this is that if you are designing an array of icons that people will have to search frequently, it is sensible to give them different pragmatic subjects, rather than giving them different predicates.

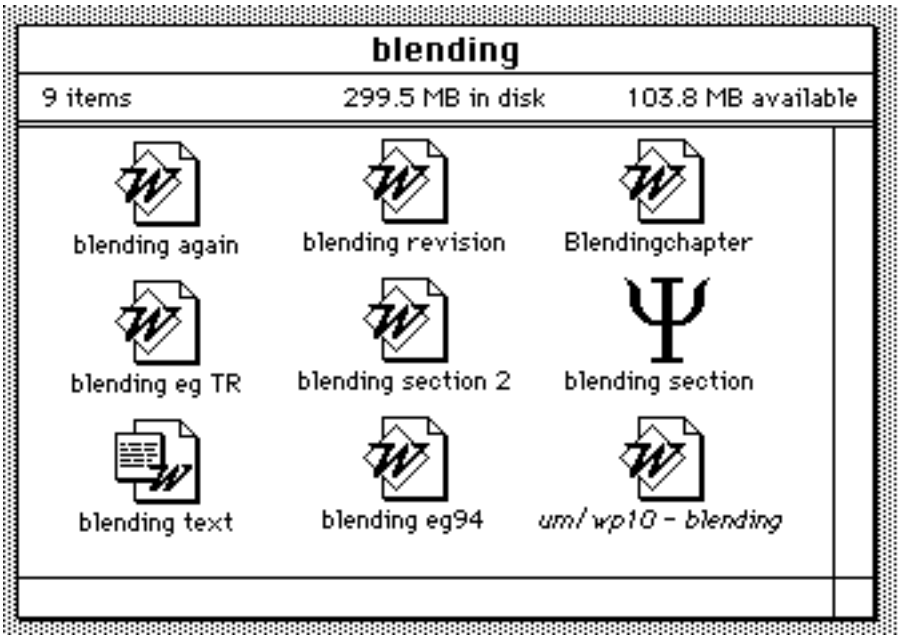
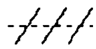


Figure 29: pop-out in an array of icons

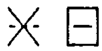
Exercise Set 4

1. Draw structure diagrams for each of these abstract icons (ignore the words):



delete line

word search



replace word

scroll right

2. Draw a structure diagram for the array of icons in Figure 29 (ignore the text labels and the ‘frame’ of the window), with the ‘document shape’ as the pragmatic subject of the icons that have one.
3. Draw another structure diagram for the icons in Figure 29, this time assuming that the ‘W” of each icon was its pragmatic subject.
4. On the basis of the two diagrams you have drawn for Figure 29, would it be easier to find the ‘blending text’ icon if it had the ‘document’ or the ‘W’ as its pragmatic subject?

An object's structure affects grouping

The examples of pop-out we have seen so far have shown that changing one attribute of an object can affect the structure of the scene, by determining which other objects it will or will not form a group with. As well as changes to attributes, changes to an object's own internal structure can also affect grouping: structure affects structure!

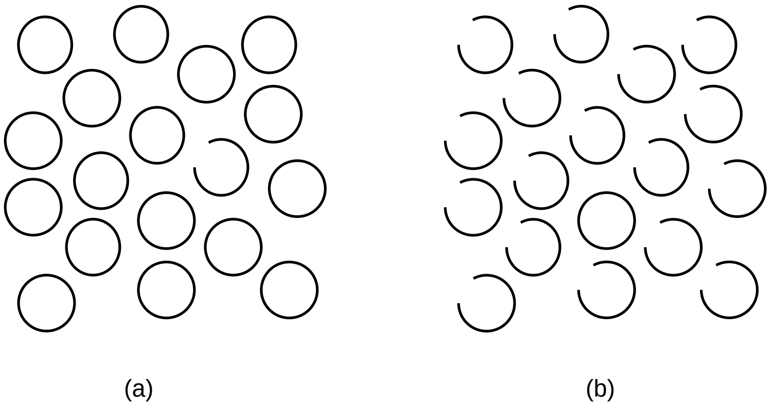


Figure 30: asymmetry of pop-out due to structural differences

The arrays (a) and (b) in Figure 30 contain two types of object. One is a simple circle, the other is an incomplete circle – a small part of the circumference is missing. When the incomplete circle is placed amongst an array of complete circles (a), it is easy to see the incomplete circle. The opposite is not true – in array (b) it is much harder to locate the complete circle amongst a number of incomplete circles.

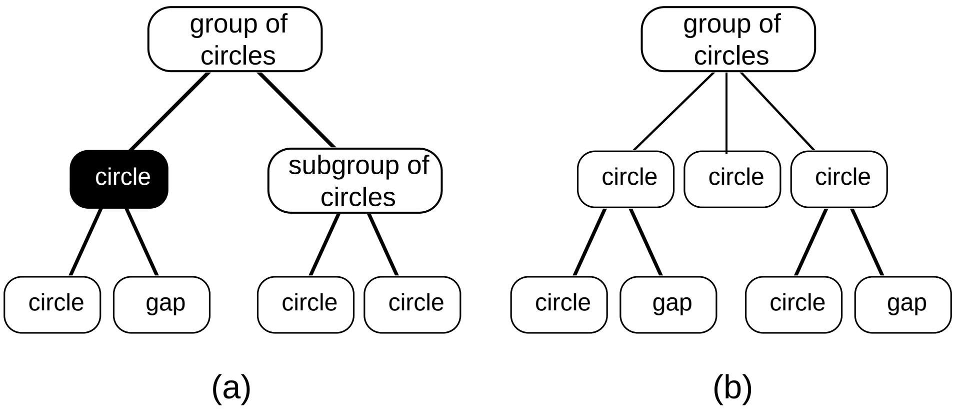


Figure 31: An object that has a structure will pop out of an array of otherwise identical objects (a), but a subject-only object will not pop-out from among more complex objects(b).

When we draw structure diagrams for these two sets of circles, we have to show the incomplete circle in set (a) forming a pragmatic subject, to make it clear that it pops-out. For set (b), we have to show the complete circle as part of the same group as the incomplete circles, because it doesn't pop-out. Figure 31 shows how these groups are composed (not all of the circles are shown).

In the figure we have described the incomplete circles as ‘circles plus gaps’ – in effect, we are saying that they actually have two components to their structure, while the complete circles are just circles, and have no further structure.

When most of the objects within a group are simple, and do not have a structure, a similar object that does have a structure cannot form a group with them at the same level. The simple objects form a subgroup, and the complex object becomes a pragmatic subject, as in 30(a). In contrast, when most of the objects within a group are complex, with a common pragmatic subject and a structure, simple objects that consist of the same pragmatic subject but nothing else, are able to join the group: as in 30(b), they are simply perceived as similar to the other objects, but less complex. They are able to ‘hide’ amongst the noise of the other objects structural complexity.

Complex objects that do have a structure pop out from simple objects that do not have a structure.

Simple objects can ‘hide’ amongst more complex ones, if they have the same pragmatic subject.

Knowledge can affect structure

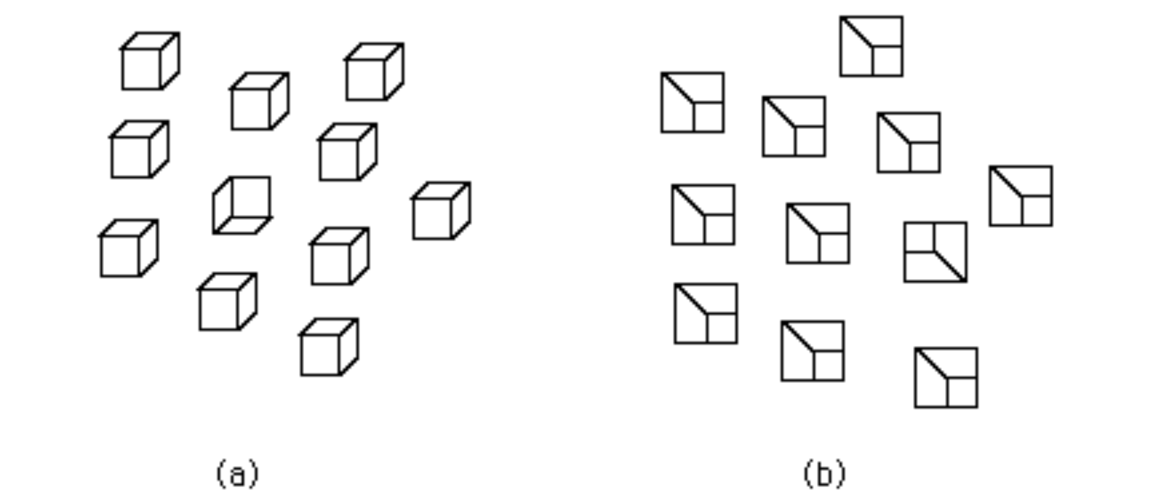


Figure 32: two very similar sets of objects with different pop-out effects

We have now seen that structural information can determine pragmatic subjects – it isn't all to do with lines and colours. Pop-out also occurs with items whose size, shape and colour are the same, but whose structures are different, as in Figure 32(a). The 'cube' that points the 'other way' stands out. You might think that this is just because it has been rotated, but there is no pop-out with similar objects that contain the same number of lines, angles, and so on, as in 32(b).

Pop-out for the objects in 32(a) must be due their grouping in the way that is shown in Figure 33. Lack of pop-out for the objects in 32(b) must be due to their forming a single group, which needs to be searched through for the 'different' object to be found.

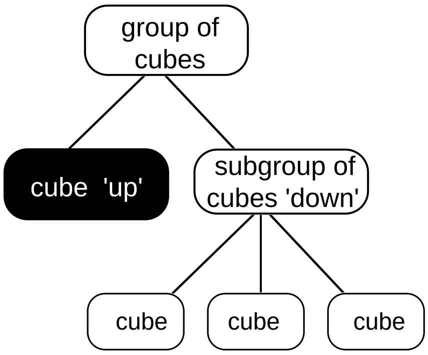


Figure 33: the structure diagram of the 3-D cubes.

This is a good example of how the knowledge that a person has access to can affect how they perceive the display. If you are told that the objects in 32(b) show the end of a megaphone, or an empty box 'descending' into the display (as in Figure 34), then you can form a mental image of a three-dimensional depth relationship between the lines on the screen. Nothing has changed visually, but now the one item 'facing' the other way pops out of the display.

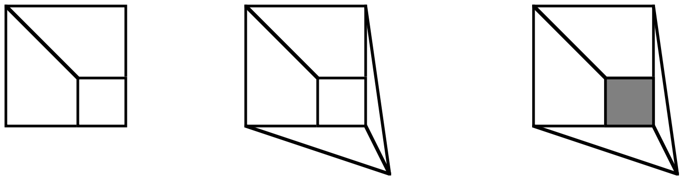


Figure 34 - reinterpreting the structure of the objects in Figure 32(b)

Seeing it as three-dimensional has made its 'direction' obvious, and it has been grouped separately from the other 'empty boxes'. You have almost certainly seen lots of pictures of cubes like those in 32(a), and so you were able to see them as three-dimensional straight away, but the use of perspective in the objects of 32(b) is unconventional, and so you had to be given a hint as to how to interpret them as three-dimensional.

Until now we have just described information that the viewer of the world has as being either a visual or an object representation, with a transformation process that turns the visual level information into the object level (this was illustrated in Figure 13). To allow knowledge to influence the object representation, we need to add another level of information. The object level was more 'abstract' and less detailed than the visual level, but contained more structure and interpretation. In the same way, the new level is more abstract and more interpreted than the object level. It represents 'semantic' or 'propositional' knowledge about objects – their names and properties, and the way that they can be expected to relate to each other. In Figure 34, the object representation might consist of one square within another, their corner linked by a diagonal line, but the propositional representation of the same drawing can be of an 'empty 3-D box'.

This propositional level of representation can be produced from the object level by an object-to-propositional transformation, which examines the object representation, identifies the objects, and produces a 'meaningful' interpretation. The visual-to-object transformation added structure to the visual representation, and the object-to-propositional transformation adds information that wasn't necessarily there at all in the visual representation. Importantly, the propositional representation can also be used to produce new object representations, by a propositional-to-object transformation. This is how the knowledge that the object in Figure 34 is 'really' an empty box can feed back to influence the structure of the object representation. This feedback process is illustrated in Figure 35, which extends Figure 13.

The 'active nature of perception' that we began this guide with is becoming much more active: with the addition of this feedback loop, it becomes possible for the object representation that a viewer forms at one moment to influence the object representation that is formed the next moment. The object representation is receiving information from both the visual-to-object transformation and the propositional-to-object transformation.

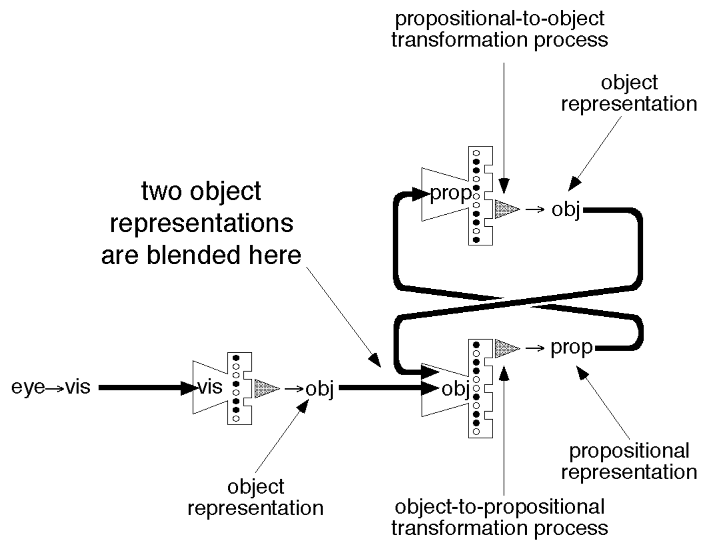


Figure 35: the feedback loop between the propositional and object levels of representation

The **object-to-propositional** transformation process identifies and relates objects

The **propositional-to-object** transformation process feeds back information about object structure

The representation that the object-to-propositional transformation actually receives is a result of these two inputs being blended together: the parts that match reinforce each other, and parts that don’t match are discarded. This can be of great benefit in perception, where the sensory, visual level of representation is often incomplete or distorted. When the visual-to-object transformation is unable to produce a clear object representation, the contribution of propositional knowledge allows the viewer’s expectations and knowledge about the world to clarify matters.

The object representation that is perceived is a **blend** of information from visual and propositional sources.

The object in Figure 36 has the identity ‘teddy bear’ for a viewer who has already learnt the propositional representation of such an item, and knows that teddy bears generally have a head with ears that stick out, and limbs that are spread out. Even if the viewer of this figure has no idea of what they are about to see, the extreme familiarity of this outline enables the feedback between the object and propositional representations to settle on this interpretation very rapidly, perhaps ignoring any visual features that did not quite ‘fit’ – the gaps in the outline, for example. People have a strong tendency to give objects nameable identities if they possibly can, reflecting the extraction of propositional knowledge. These names then affect the way the objects are perceived. The propositional influence on the perception of familiar forms like this is very resistant to distortions in the shape, provided that key invariants between the objects are met. If Figure 36 is looked at upside-down, for example, a different teddy bear can be perceived (or perhaps the longer ‘ears’ now make it look like a cuddly rabbit).

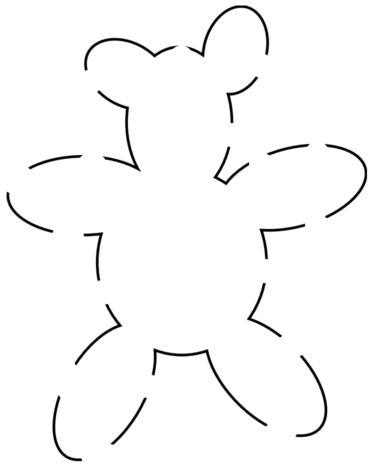


Figure 36: propositional knowledge helps to 'complete' the object representation of this figure

For someone who has never seen a teddy bear, and who has no other propositional representation that 'leaps in' to influence the perception of a single object, the form may appear to be a number of overlapping circles and ovals, but the absence of arcs in the centre of the form completing or even continuing these circles means that it is much more likely to be seen as a single, irregular shape.

Propositional representations can be used to fill in gaps in object representations that have been derived from incomplete visual representations.

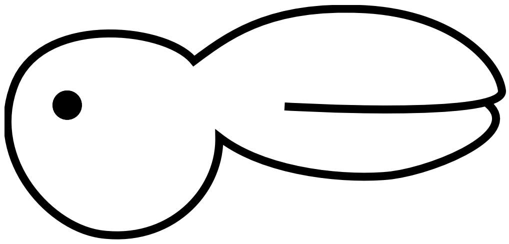


Figure 37: ambiguous form

The feedback loop between propositional and object representations tries to settle on one consistent interpretation of a figure. The form in Figure 37 can be seen as either a rabbit (looking to the left) or a duck (looking to the right), but it cannot be seen as both at the same time: the perception must 'reverse' between the two interpretations. Notice that the propositional identity given to the figure constrains the structure of the object representation – the beak becomes a pair of ears, and the direction that the eye is looking changes. As with the 'empty boxes' in Figure 34, these structural changes in an object representation that has been derived from a single visual representation are indicative of propositional knowledge being brought to bear.

Once an object has been propositionally identified, we are able to 'go beyond' the available sensory data to use our knowledge about the world to enhance its object representation. If we are told that an object is 'round', or has a 'hole', then we can combine the sensory information that is available now with information that we have experienced in the past as being common to ‘round’ or ‘holed’ objects (Figure 38). If we were told that it was ‘round’, we might actually interpret it as ‘spherical’, even if the appropriate sensory information (such as shading) is not immediately available. If we were told that it was a ‘hole’, we might be able to perceive some visual features as belonging to another object that is visible through it.

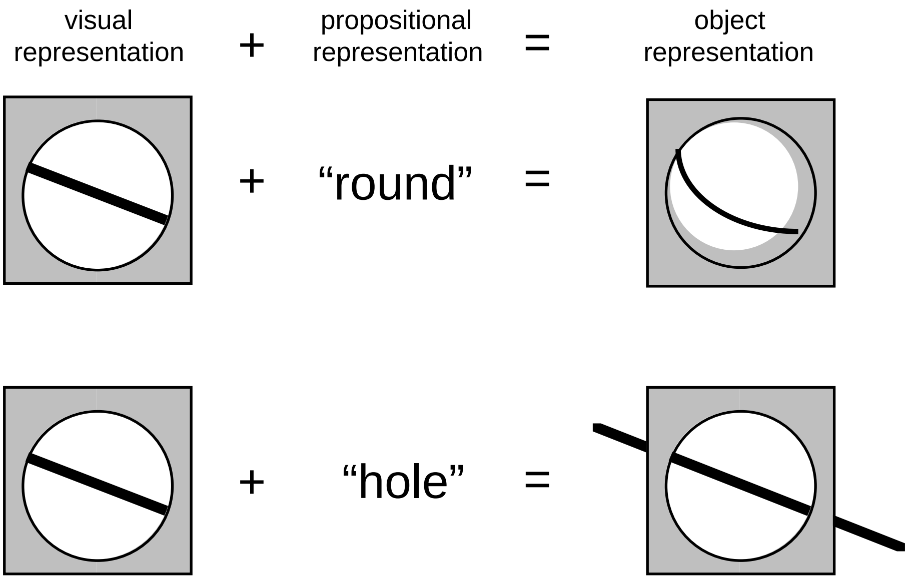


Figure 38: the same visual representation can result in different object representations, depending upon the contribution of the propositional representation

Propositional representations help the object representation settle on one interpretation of an ambiguous figure. The perception of ambiguous figures depends on what the viewer knows, and what they expect to see.

Exercise Set 5

1. Draw the structure diagram for the part b of Figure 32, showing the representations formed without any propositional knowledge, so that objects are perceived as two dimensional.

2 Now draw the structures with propositional input, so that they are perceived as 'empty boxes'. Which object is the pragmatic subject?

3. Look back to the structure diagrams that you drew for questions 2 and 3 in Exercise Set 4 (for the document icons). Which of the document icons is the most 'complex'?

4. Draw structure diagrams for these two arrays of icons, and identify which icon, if any, is the pragmatic subject:

Learning the meaning of objects

A computer user who was searching through the abstract and representational icon arrays of Figure 29 had to generate mental images of 'target' icons that they then compared the 'candidate icons against. Before the experiment could begin, the users had to learn what each of the icons 'meant' – this is propositional knowledge. When they were then asked to 'find the icon for *delete line*', they were able to retrieve their propositional knowledge and use it to generate an object representation of the icon that they were looking for, without any visual information.

In practice, users can never be given exhaustive training on computer programs, to make sure that they know exactly what every icon looks like and means, nor what the structure of every dialogue box is going to be. What users will learn about the structure of interface objects depends on how they use them, as was shown by two hypertext databases that used versions of the same display design, with one slight difference (Figure 39).

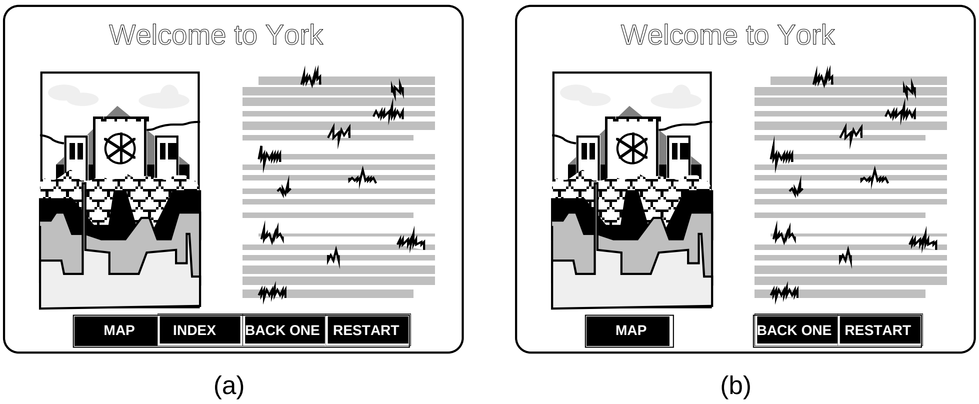


Figure 39: The two versions of the York Hypertext interface

The prototype 'visitors guide' had been built to let people read text and see pictures of York. They could navigate by clicking on 'hot spots' in the text, or by using a set of buttons at the foot of the screen. In Version A of the interface, these buttons allowed them to access a schematic 'map' of all of the screens, an alphabetic 'index' of the screens' titles, to go 'back-one' to the previous screen they had seen (rather like an undo function), or 'restart' to go right back to the first screen.

Version B was the same, except that the 'index' button was omitted. People were shown how all of the functions worked, and then given some questions about York that they had to find the answers for in the hypertext. A typical pattern of exploration involved the users selecting a screen, and then realising that it didn't offer them any help, and so they would use the 'back-one' button to retrace their steps. Sometimes they would get completely lost, and 'restart'.

In Version A the navigation buttons were presented together as a block at the bottom of the screen, and since they had similar shapes and colour, they formed a group on the screen (Figure 40). To attend to any one of these buttons, people first had to attend to the group as a whole, and then make a transition into the group's structure.

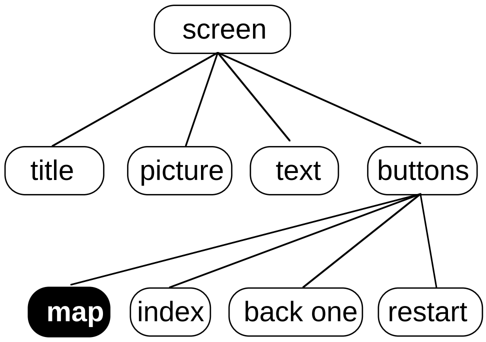


Figure 40: the structure of Version A, with the ‘map’ button forming the subject following a transition into the structure of the navigation buttons

When they did ‘zoom in’ to the structure of the groups, the individual objects were the four buttons, but since they were all identical (apart from the textual labels), none of them ‘popped out’ as a pragmatic subject according to shape or colour. In these circumstances, users would be most likely to scan across the buttons from left to right, reading the labels as they would normal text. A transition path that would be required to find the ‘restart’ button in this interface is shown in Figure 41.

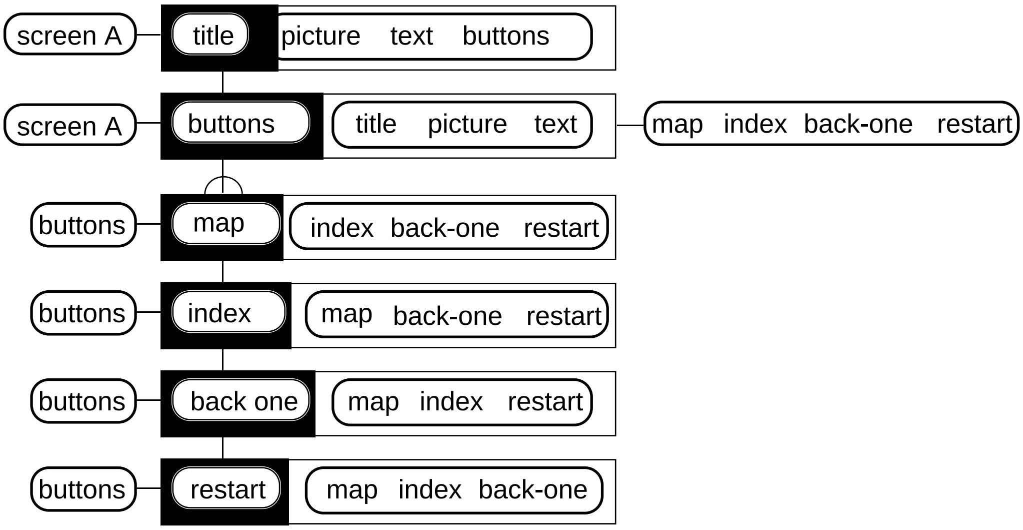


Figure 41: transitions required when using the buttons in Version A

After people had answered all of the test questions, they were asked about the various functions and buttons, and what they all did. Most of the people who had used Version A were found to understand that the map and index buttons could help them navigate around the system, even if they had not actually used them (they had, after all, been shown them in the introduction). Surprisingly, the people who had used version B were found to have less knowledge about the purpose and use of the ‘map’ button. A look at the structure of the interface shows why this might be (Figure 42).

In Version B, the gap left by the omission of the ‘index’ button breaks up the group of navigation buttons. Now the ‘map’ button stands on its own - and depending upon the size of the picture, might actually be associated structurally with it instead of with the other buttons. When people had to use ‘back one’ or ‘restart’, they no longer had to encounter the 'map' button (the transition path is shown in Figure 43). This meant that as they used the system, they did not encounter the 'map' button while they were navigating, and so the information they had been given at the start of the session about its function was not integrated into their propositional understanding of the system's functionality.

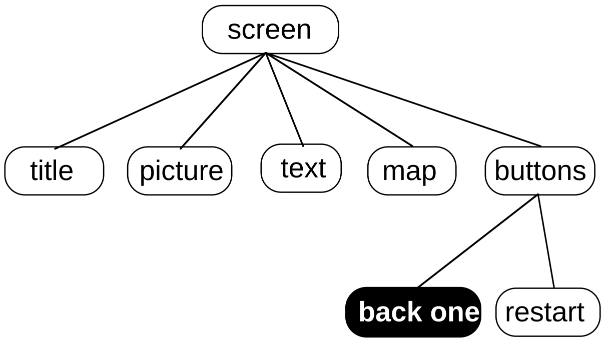


Figure 42: the structure of Version B, with the 'back one' button forming the subject following a transition into the structure of the navigation buttons.

The transition path diagram for Version B shows that it is indeed easier for the users to find and use the back-one and restart buttons, but it is at the cost of their understanding of the map button. A conclusion we can draw from this example is that grouping items together visually not only helps users to locate them *as a group* when they need to use them, but it also helps them to generalise about their common functionality. Propositional knowledge about one of the group will tend to 'rub off' on the others. Of course, this can only be beneficial to users if there really is some similarity in the functionality of the grouped buttons. If the 'map' button had actually shown a geographical map of York, rather than of the hypertext, it would not have helped them to navigate around the system at all, and so it would have been misleading to place it with the other navigation buttons, even though they are all 'objects the user can press'. Functionality must be thought of in terms of 'what the user wants to achieve' rather than 'what the system lets the user do'.

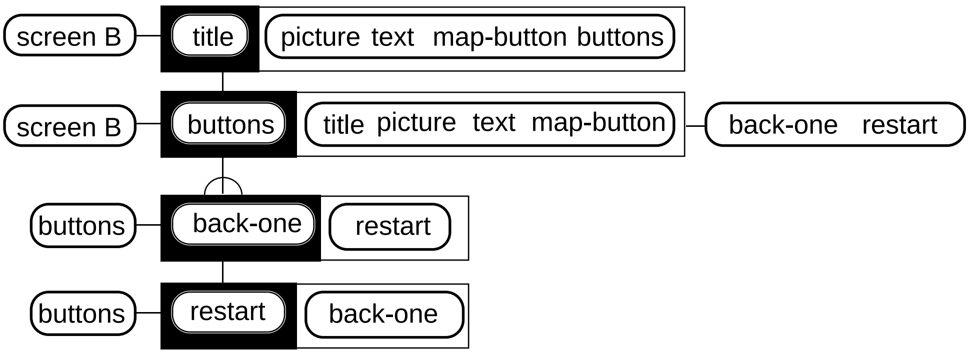


Figure 43: visual transitions when using buttons in version B

The structure of objects affects the way they are used, and so affects the propositional knowledge users learn

Competing groupings

The contribution of propositional knowledge to perception might seem to make the task of display design a whole lot easier: after all, if users can be told what to look for, and can learn how to group objects, why should it matter what the visual information is like? The pop-out examples shown earlier in this guide should convince you why this argument fails, because even when you know that all of the circles are just circles, the differently coloured circle still pops-out.

Those examples were selected especially because they showed the effects of pop-out very strongly, of course. They were generally very simple, with only one attribute changing to influence grouping. Even in more complex arrangements, the dominance of the visual contribution to object representations can be just as convincing.

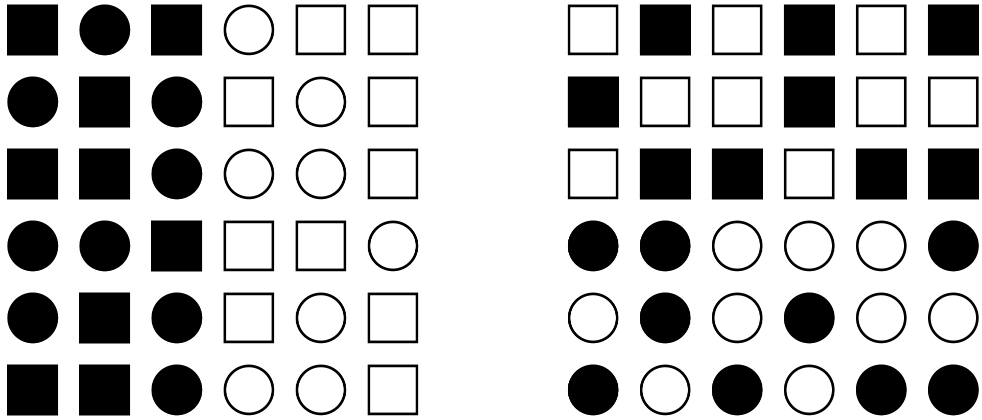


Figure 44: Colour dominates shape

The left hand array in Figure 44 shows again how strongly colour can determine the grouping of objects. It is easy to see this as a group of black objects and a group of white objects, even though the objects themselves have different shapes. In the right hand array, where the objects are arranged by shape, it is not so easy to see two distinct groups.

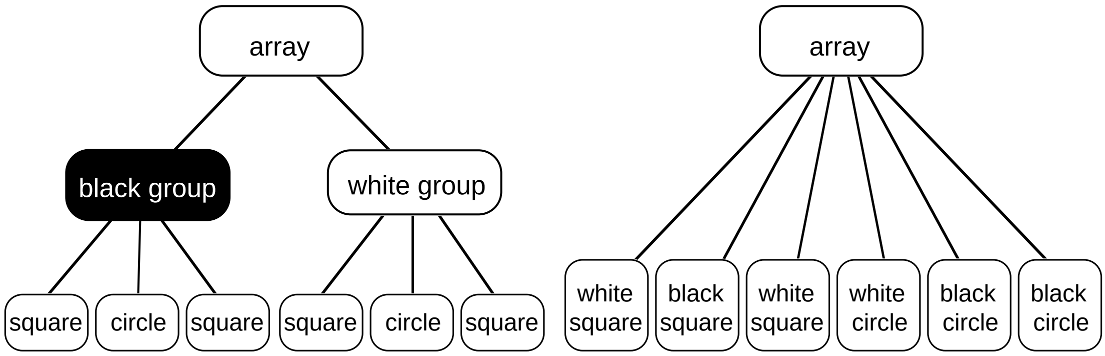


Figure 45: the structure diagrams of Figure 44

Here the randomness of the colour dominates the orderliness of shape, preventing the visual-to-object transformation forming a pragmatic subject in the structure of the array. Figure 45 shows structure diagrams of the two arrays. In the left hand array, there is an ‘intermediate’ level of grouping between the array and the individual objects, but this is missing in the right hand array. Once you ‘know’, propositionally, that there are two ‘shape’ groups in the right hand array, you can impose this structure on the object representation, but it seems to require continual mental effort to do so. As soon as you look away and back, the randomness of the colours dominates once again.

In grouping objects, **colour dominates shape**

The arrays (a) and (b) in Figure 46 show pop-out due to the shape of the oblongs. In both (a) and (b) the object in the array that does not have the same shape as the others becomes the pragmatic subject and pops out. In (c) the object with the differing shape also has a different colour, and the pop-out effect is enhanced. As in Figure 44, colour is stronger than shape, for in (d) the white, horizontal object pops-out, while the black, vertical form does not. This is despite the fact that the white object has the same shape as the black objects, while the black object has a different shape: so colour dominates shape.

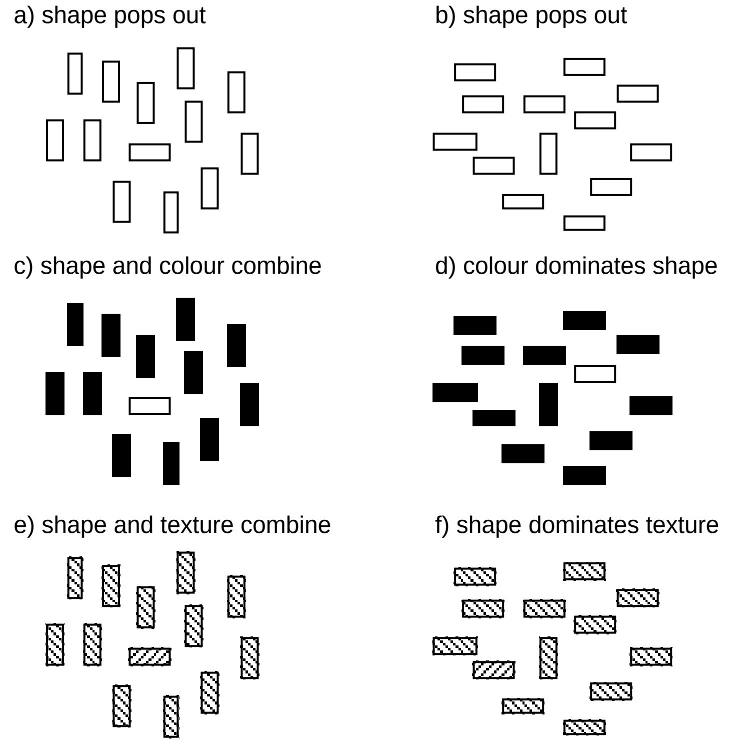


Figure 46: varying combinations of colour, shape and texture

Arrays (e) and (f) introduce a new attribute, texture. Texture is defined as a regularly spaced, repetitive pattern where each of the objects of the pattern is individually perceptible, but where the objects are much smaller than the whole that they fill. Although it is common to think of texture and colour as very similar, since they appear to be properties of the surface of objects, visually they have quite different properties. Unlike colour, texture does not dominate shape.

In (e) the horizontal oblong with a different texture (direction of stripes) to the other oblongs stands out from the array, but in (f), the upright oblong that has the same direction stripes as the other oblongs stands out instead. The horizontal, differently textured oblong is still in the array, but it is much harder to locate. Clearly the 'difference' of the texture is not as influential as that of colour. Like size, brightness and contrast, texture is also a 'depth cue', with objects of a finer texture appearing further away than objects of a coarse texture (Figure 47).

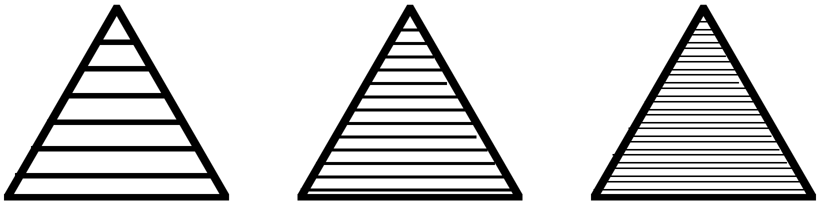


Figure 47: objects with coarse texture appear closer than objects with fine texture.

Although the textures in Figure 46 were differently sloped, they were equally bright, and so they did not suggest that any of the oblongs were any nearer or further away than any of the others. This meant that the texture attributes did not force them to join separate groups, allowing the shape to dominate.

A coarse texture, like a larger size and increased brightness, makes an object look **near.**

In grouping objects, colour and shape ('difference') dominate nearness.

The grouping of objects on the basis of attributes like colour, size and shape is such a pervasive part of design that we tend to take it for granted in many circumstances. Text is a good example. Letters that are written in the same font and face are easily grouped into words, even if they do not actually spell recognisable words. Changing the attributes of letters makes the words much harder to read, because their component objects become more 'visible' than the whole word (Figure 48), even though the word boundaries (spaces) are still there.

letters in similar faces form words

digghe huyttr noggel guty ssinf buttr

dissimilar letters are harder to read

Figure 48: similar attributes of letters help them form easily readable words, even if they are nonsense, but letters with dissimilar attributes are harder to group into words, even if they make sense

Exercise Set 6

1. Draw a structure diagram for parts c, d, e and f of Figure 46.
2. Identify the pragmatic subject of the main group in each diagram.
3. Which of the subgroups also have pragmatic subjects?
4. According to the answers that you have given for question 3, is it easier to locate the differently shaped oblong in part d, or the differently textured icon in part f?

Icons for multivariate information

Some designers have tried to make use of our skills in the recognition and identification of complex graphical forms to represent complex 'multivariate' data with icons. Multivariate data is obtained when things are measured on several variables at the same time. For example, a house can be measured on its price, number of bedrooms, distance from the station, and so on. These measurements could be shown quite concisely in a numerical table, with a row for each house, and a column for each measurement, but trends that involve more than one variable are hard to detect from numbers alone. House price may be directly related to the number of bedrooms, or the distance from the station, for example.

Graphs that plot one measurement on the horizontal axis and another on the vertical axis are better, since the spatial groupings and positions within the area of the graph directly show the relationships between two variables. Even four measurements require six graphs to show all of the pairwise relationships: and they cannot show relationships that involve more than three variables at all.

The icon solution tries to represent each measurement by a different attribute of an object, and then rely on the viewers' ability to use propositional knowledge to concentrate on the relevant parts. In Figure 49, each of the 'faces' has three attributes that can change: the size of the nose, the curve of the mouth, and the angle of the eyes can each vary independently. If the nose represented a house's price, the mouth its number of bedrooms, and the eyes its distance from the station, the viewer of this array could use their propositional knowledge to generate a mental image (an object representation) of, for example, a medium nosed, flat mouthed, slant-eyed face, and then search the array for that icon.

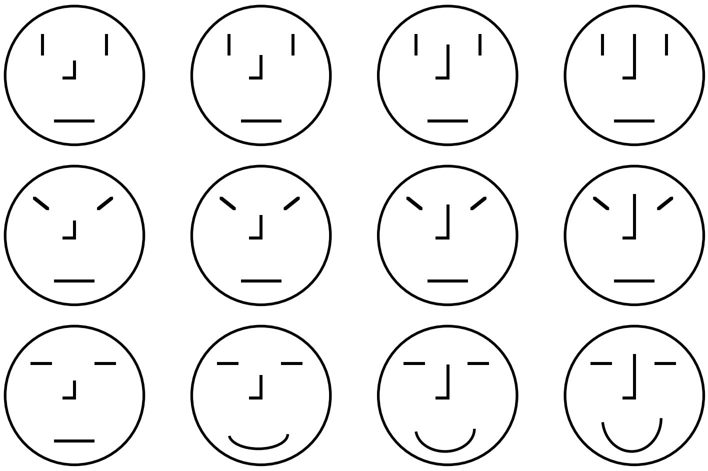


Figure 49: multivariate data presentation using faces

In practice this system has proved hard to use. If we look at the structure of these faces, they resemble the control-panel icons of Figure 27 and the representational word-processing icons of Figure 22: they all have the same general 'shape' and border. The faces therefore form an array, but no subgroups, and require search within the structure of each face to get information about how the faces differ, and therefore about the variables that the users are required to judge. As with the control-panels and word-processing icons, this predicate search requires transitions up and down within the structure of each icon, as well as transitions between the faces, which slows search down. The search and comparison of the face icons is represented in Figure 50.

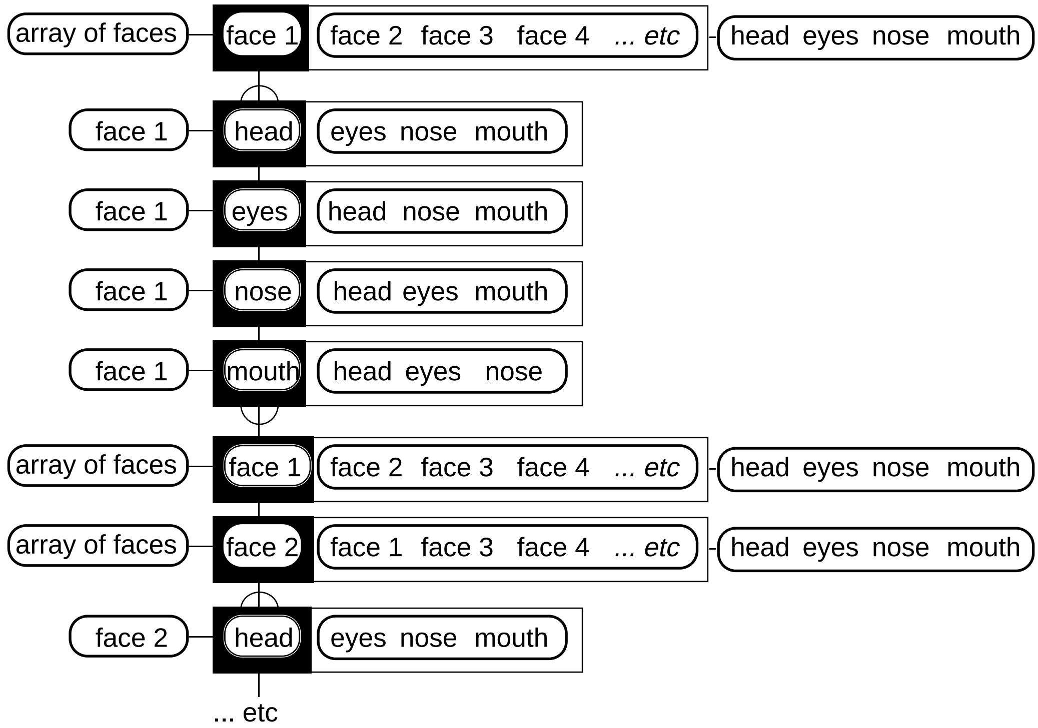


Figure 50: the transition path required to make decisions about the attributes of a multivariate face icon

This is an example of the visually derived object representation being too strong for the propositionally derived knowledge to influence. With practice, users of the face icons can certainly generate a target image to search for, but it simply isn't possible for them to group the faces according to any particular attribute or combination of attributes.

A visually derived pragmatic subject can prevent the propositional representation influencing structure

A second family of icons to present multivariate data uses the structures of stick men (Figure 51). In this system the different dimensions are represented by varying the head size, body size and slant of limbs. We have shown that attributes like these can affect grouping in very simple visual items. In experiments, this system of representation is easier for subjects to use.

Because these stick men have no pragmatic subject, the propositional knowledge can now influence the object representation. Any one of the different attributes could form the psychological subject. In practice, this means that a stick-man (or a group of stick-men) with a different value to its neighbours for any one of its attributes can be made to pop-out, and be easily identifiable, provided that the user of the array is ‘concentrating on’ that attribute. Furthermore, an object representation of the target will now have as its subject the same attribute that has been used to group the objects in the array. This representation can then be used to drive the search, rather than letting it be constrained by the visual features of the array. This reduces the need for predicate depth search and shortens search time.

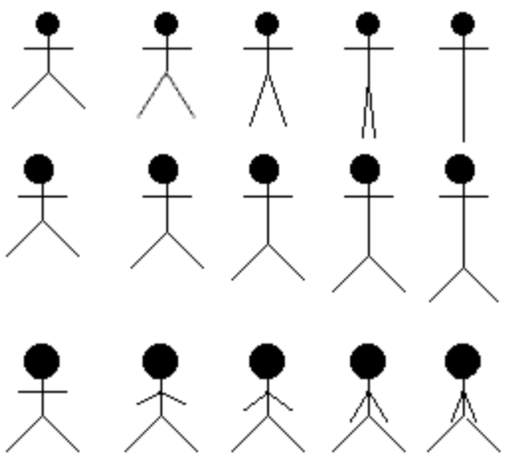


Figure 51: multivariate data presentation using stick men

If propositional representations can be used to group objects as well as to generate a target to search for, search is easier and quicker.

Dynamic changes in structure

The examples so far have concentrated on static screen displays – ones in which nothing moves. Clearly, motion affects the structural attributes of displays: one moving object against a static background is very likely to pop out, and grab people's attention, regardless of its other attributes. If many objects move in an unco-ordinated fashion, the result is just going to be confusing, and very hard to make sense of. Between these extremes you might be able to see how the principles of grouping that we have presented for static attributes can also apply to dynamic attributes. Consider a group of people walking along a street. Unless they are an army marching in step, they will all move at slightly different speeds, and yet we can perceive them as a group, because they all move in *roughly* the same direction, at *roughly* the same general speed. Similarly, a few screen objects with similar attributes of motion can be perceived as a single 'group' moving on the screen.

Motion is an extremely salient visual attribute, and can even dominate colour and shape.

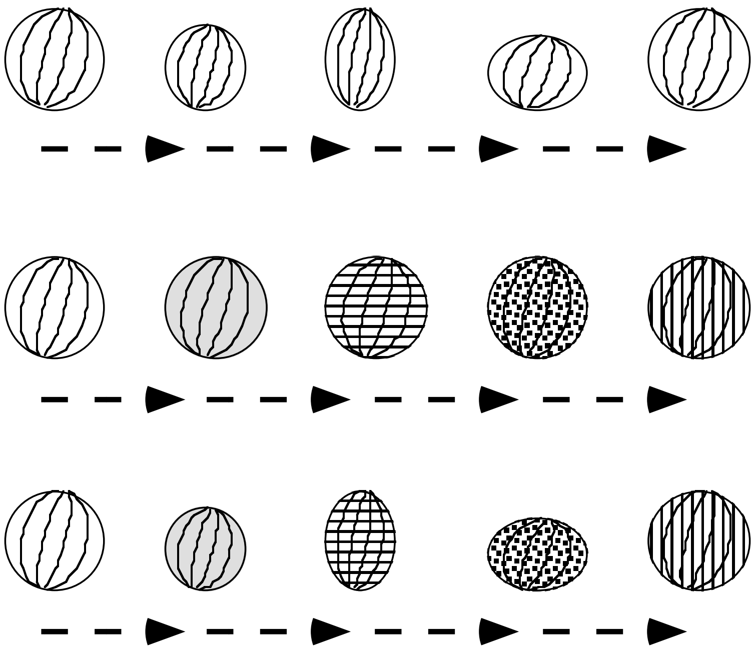


Figure 52: a dynamically changing ball can be seen as such (top two rows) if the changes over time are not too great (lower row).

Other attributes can also vary over time - an object might change colour, size, or shape, for example. This is a key principle in animation, and allows us to see differently drawn views of an object as the 'same thing' changing, rather than as 'different things', replacing each other. For this to work perceptually smoothly, the viewer must be able to construct a smoothly changing object representation, and so the visual changes must not be too great. In the top row of Figure 52, a ball changes shape over time. If this were an animation, we would easily see this as a single, changing ball, rather than as separate balls replacing each other. In the middle row, its colour changes but its shape remains constant. Again, it is easy for us to see it as a single, changing ball.

In the lower row, both of these attributes change at the same time. Now it is harder (although still possible) for us to see it as a single ball. This sequence of changes, if animated, appears much less ‘smooth’, and it feels as if we are seeing several different balls. This is because the overall change in the attributes of the ball has become too great for the visual-to-object transformation to produce an object representation that can coherently blend with ongoing processing, in particular the product of the propositional-to-object transformation, which was based on the object representation of the previous ball.

Objects that change over time can be perceived as the same object, provided that the visual changes are not too great to be reconciled with the propositional representation

Animation is a familiar example of dynamic, changing drawn or computer generated objects. Ordinary films also tell us something about the management of dynamic changes in display design. Over the last century, film-makers have developed editing techniques that allow them to cut from camera to camera, dramatically changing the structural contents of the visual scene, without confusing or misleading the viewers of their films. They can jump spatially between different viewpoints within a scene, or temporally, skipping over periods when nothing interesting is happening. They can even intercut different scenes without confusing us.

Many of the rules of thumb that film-makers follow when cutting films together have to do with the content of the narrative, but others have to do with structural details of the shots either side of the cut. When a film-maker cuts from a view of someone firing a gun, to a view of his victim falling, a conventional cut will place the victim in roughly the same screen location as the gun (Version A of Figure 53).

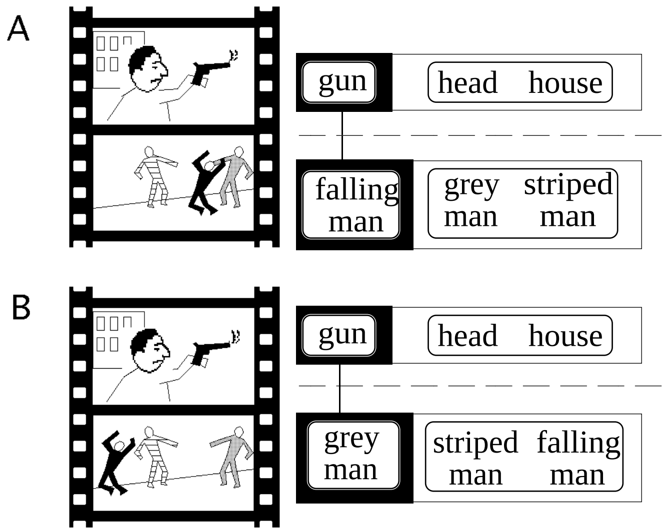


Figure 53: successful cutting in films is dependent upon thematic continuity in the structures.

Before the cut, viewers will have been watching the man raise the gun, and they are likely to have had the gun as the psychological subject of their object representation. When the cut occurs, they will be looking directly at the falling man, and so this will form the immediate pragmatic subject. Since this fits coherently with their ongoing propositional comprehension of the scene, providing ‘thematic continuity’, the cut makes sense and seems perceptually smooth.

Version B of Figure 53 shows an unconventional cut. The falling man is not ‘collocated’ with the gun, and so following the cut, the viewer has to search the scene to find an object that makes propositional sense. This cut feels less perceptually smooth. This type of cut is more likely to be ‘noticed’ as a perceptual ‘jump’ because it does not provide thematic continuity.

A similar use of collocation is made when film-makers zoom-in to a scene to provide greater detail on some object. Here the rule is that the same object should be the subject of the before and after shots, and that it should be collocated with itself. Clearly, to do this the film-maker needs to ‘know’ what it is that the viewer is going to be looking at, and so they will often try to direct the viewers’ attention towards the object that they are going to zoom-in to. An actor might pick the object up, or direct their gaze towards something, so that the viewers also look at it.

Thematic continuity, through collocation of objects that are visually or propositionally related, helps users orient themselves over screen changes.

Many applications need to change screen displays, so that all of the information changes. The advice of film-makers would be to provide some thematic continuity over the screen change, just as they do over cuts. An example of this sort of design problem can be seen in Figure 54, where a tourist information system was being implemented on a PC. The screen had to display a large scale overview of the whole area (the left screen), and allow people to ‘zoom-in’ by clicking on a particular place (right screen)

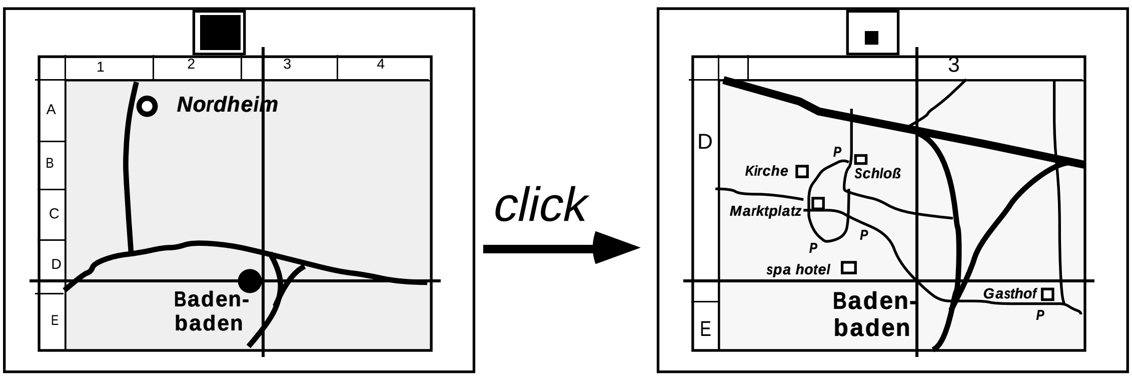


Figure 54: when the user clicks to zoom-in, the display provides thematic continuity

Just before the user clicks, they will have moved the cross-hair to the area they are interested in, and will be attending to that part of the screen (perhaps to the name of the city, to the blob marking its position, or to the road layout). After the screen zooms-in, the display preserves some of these objects, and collocates them. If the user had been attending to the name or to the road junction, then these objects would still form the psychological subject, even though their size has increased (Figure 55). This would give a sensation of ‘getting closer’, as well as providing thematic continuity, so that they would not have to search around the screen to find out what had happened. The ‘city blob’ has disappeared, of course, so if they had been attending to this they might be a little less sure of what had happened, but several objects of the predicate would remain (the name and the junction), and these too could provide some thematic continuity.

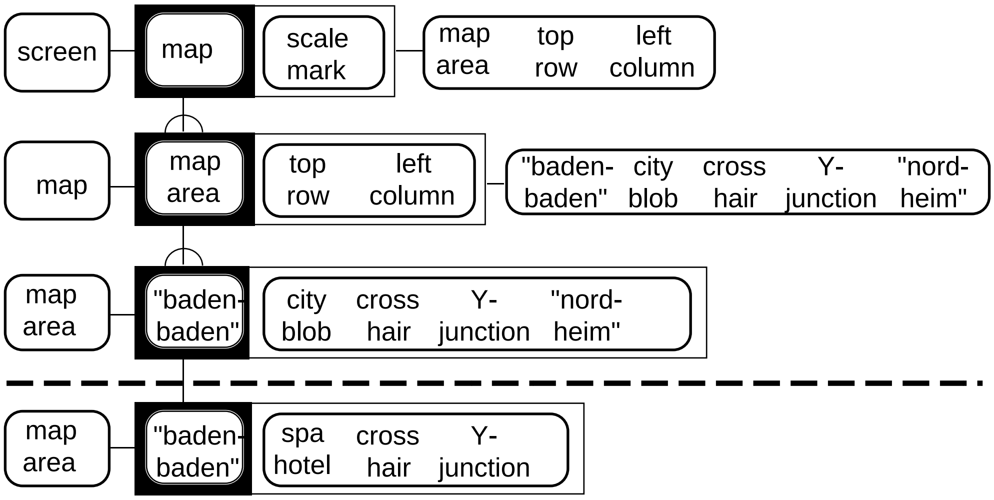


Figure 55: a transition path across a screen change (dashed line) in Figure 48 shows how the subject is maintained

Complex multiwindow displays

The interface shown in Figure 56 is a system that lets travellers make enquiries about internal flights in America. Travellers can use a mouse and keyboard to enter information, and can also input speech by holding down a button (The microphone icon in the 'Record' window). The translated speech input is shown in the 'Recognition' window. The multimedia aspects of this interface are not at issue here. We are going to consider the visual structure of the interface and think how this might affect the user's tasks.

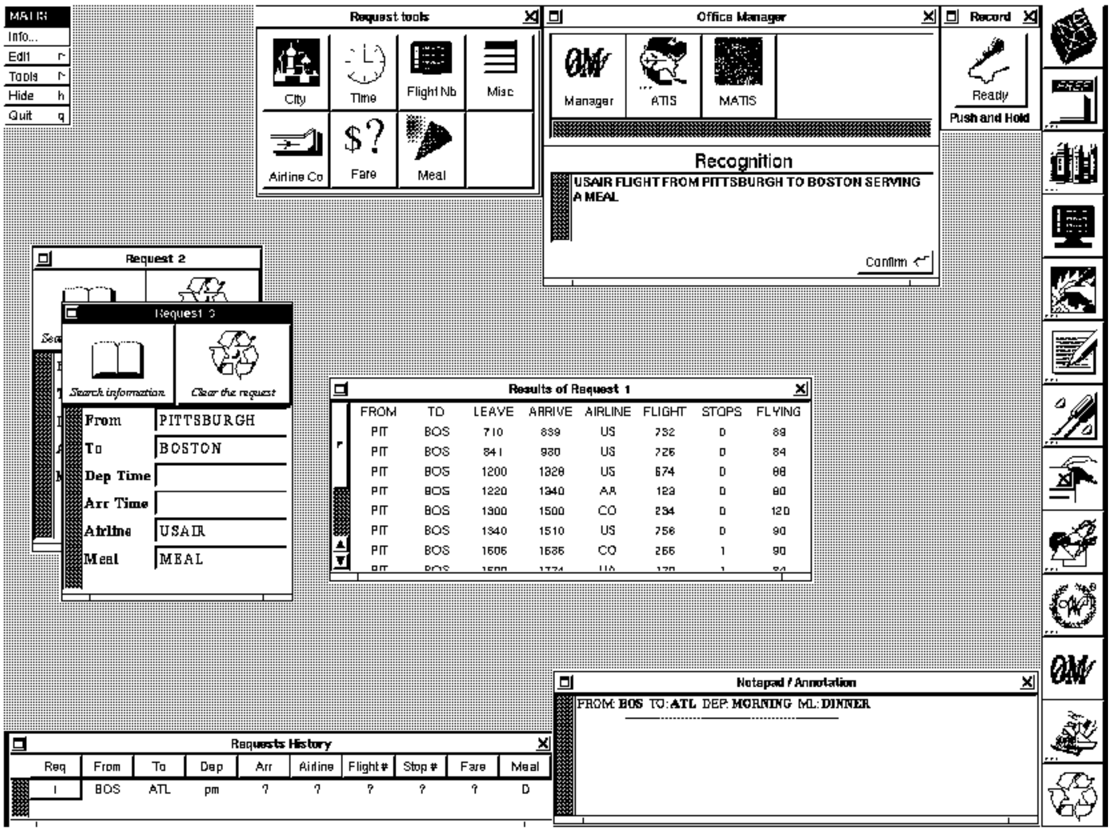


Figure 56: the MATIS screen display

Figure 57 shows the first level of a structural diagram for this screen. Because all of the windows can be moved around, resized, and repositioned by the user, we can't really tell anything about the groupings that they might form in practice (in Figure 56, for example, there is a cluster at the upper right, a cluster in the centre, and a group along the bottom of the screen). The arrangement of the windows within the screen has been maintained in the structure diagram, to help you identify the screen objects that they refer to. This can be a useful technique for complicated displays, but it only really works for showing one level of a single object's structure at a time.

If we just look for now at the requests, which are on the left of the screen, we can produce a structure diagram of one of them (Figure 58). In this diagram we have included several levels of structure, and so have not been able to maintain the spatial organisation of the objects, as we could in Figure 57. We have been able to indicate which object within each level, if any, forms a pragmatic subject. The window consists of the heading (which contains the name of the request and a close box), the icons (one to start a search, the other to clear the form), and the request form itself. This has a scrollbar and a list of search criteria, that each have a title and a slot, which will all be empty when a new request is created, but which will be filled in by the user. In this figure we've just shown the structure of the 'From' and 'Arr Time' slots, but the others have similar structures.

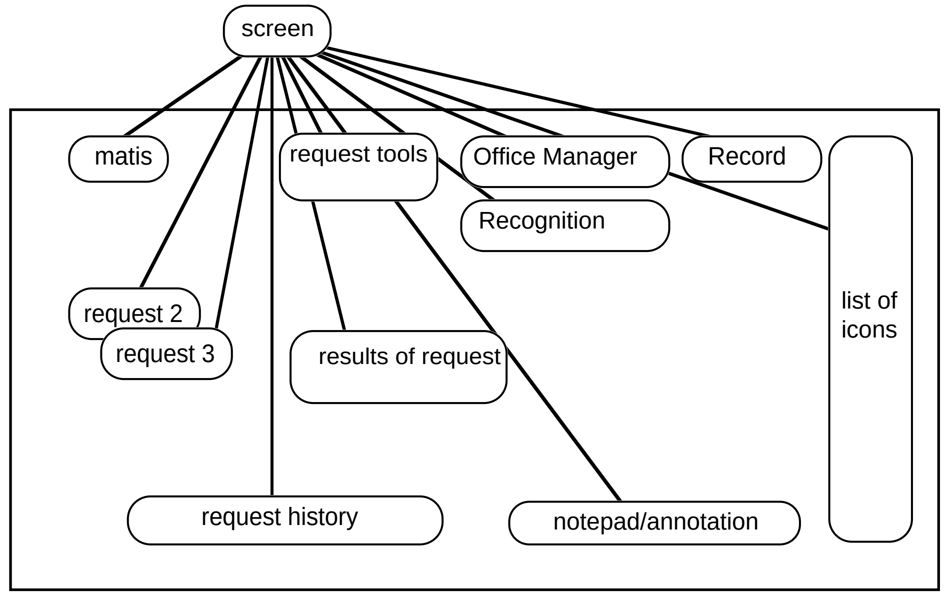


Figure 57: the top-level structure of MATIS, treating each window as a unit within the screen.

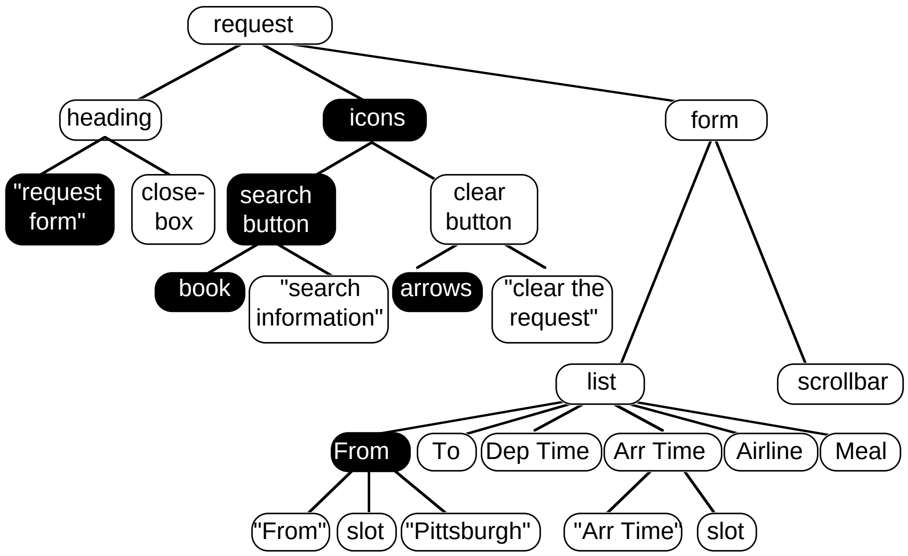


Figure 58: the structure of the MATIS request form.

What can you tell about the use of this structure? To begin with, the icons are likely to form the pragmatic subject of the form's structure, since as a group they pop out from the textual content of the rest of the objects. This is good, because in searching for this particular form, this group discriminates it from all other windows within the display – none of them contain this group or the two objects it is composed of. The user will be able to form the mental image of, say, 'a book', and will be able to reject any other windows as soon as they look at them.

Once they have found this window, they will probably want to find the slot that they have to enter a specific piece of information in. Most requests will be to find a flight from one city to another. These two slots are the ones that are most likely to be filled first on a form. The names of these two slots appear at the top of the list on the left of the window, and since people usually read from left to right and from the top of a column down, we can suggest that the 'From' slot will be the pragmatic subject of this group, and so the easiest one to find – consistent with the users' most frequent task.

The 'granularity' that you need to use to describe a structure depends upon the way that people will have to use it. As we have seen with icons, the structure of objects might need to be considered if they have to be discriminated from one another, but it may be sufficient to describe, for example, three lines just as 'lines' without further decomposing them into 'line + line + line'. In Figure 58, you'll see that we haven't decomposed words into their constituent letters, nor the 'book' and 'arrows' icons into their structures, since in both cases the pragmatic subject of each of them would be sufficient to discriminate it from the other objects within its group. If you didn't know what these icons meant, though, and were searching for a textual label that meant 'start a search', you might have difficulty locating it, since it is part of the predicate of an icon, whose pragmatic subject (a book) is not usually associated with the task of searching.

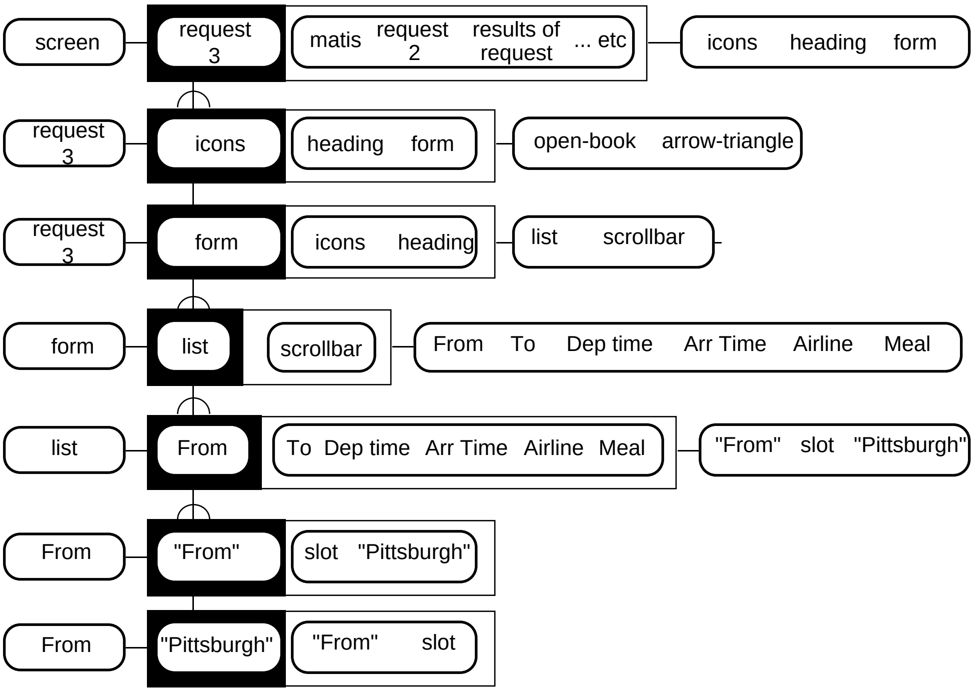


Figure 59: a transition path for the use of the request form

Figure 59 shows a transition path that a user might make to fill in a request form (that is already open), to enter the information about the city they want to fly from. We can tell from the structural description that people need to read down the list of slot-names, and then make a transition to the slot. One thing that might make it easier for this last transition to be made is if the slot names were right justified, rather than left justified, but this in turn might break up the visual structure of the list, and make the name harder to find in the first place.

Looking at the items within the list, you can also see that in an empty form, they all have the same structure. As soon as they get filled, their structure becomes more complex - an additional part is added. As with the ‘circles’ and ‘circles plus gaps’, the filled slots stand out from the empty slots, and so are more noticeable.

Exercise Set 7

1 Draw structure diagrams for the ‘Request Tools’ and ‘Results of request’ windows shown in Figure 56 – but don’t go into too much detail.

2. Complete the structure diagram of the Requests window by adding the contents of each filled and empty slot.

3 Suppose the user had just entered ‘Pittsburgh’ into the ‘from’ slot of the Request form, and ‘Boston’ into the ‘to’ slot, so that they are now looking at the word ‘Boston’ (and the rest of the slots are empty). Draw a transition path diagram to show how they would:
a) locate the search button to carry out the request
b) find a flight that departs before midday from the ‘results of request’ window (which will be the one shown in Figure 56).

4 Consider the goals the user will have at each moment in this transition path diagram(that is, before making each transition), and so what mental images will be guiding their search.
Can you spot a simple change that you would recommend to improve this display?

5. Suppose all of the slots had been filled in – does this change the way the ‘list’ object is structured? Draw a new structure diagram for the ‘list’ object.

Summary

This Guide has provided a basic overview of visual perception, in terms of movement of attention within visual structure, and has shown how the attributes of objects interact with knowledge to affect the way that people perceive their visual environment. It has touched on the effect of prior knowledge and expectations on perception, and shown how different types of computer based task can both depend upon and influence the development of knowledge. We have tried to show you that an understanding of how people perceive objects can help you design more usable displays.

In an icon search task, the more structurally similar the icons are, the better they group together as an array, and the easier the array as a whole will be to locate on the screen, but the longer it takes to search for a particular icon within the array. The more visually dissimilar the icons, the less well they group together as an array, but they can be searched faster. If an icon has a visually salient attribute that can form a pragmatic subject, this will help users to locate it.

The use of icons to display multivariate data is constrained by the grouping of the attributes that represent each variable. Some combinations of objects will be harder to 'decompose' than others, and so harder for users to search through. Conversely, searches that depend upon values of two or more variables would be easier if the features could be integrated into a single object that could be used as the subject of the search.

In the Hypertext example, the task that the user was performing (searching freely through the hypertext) was under their own control, but the display structure still affected their performance and comprehension. As they attended to the 'map' button while thinking about navigating, they came to develop a propositional representation that included navigation as one of its attributes.

In the zoom-in map, the film-makers' principle of collocation of subject over screen changes was shown to simplify the visual transitions required, and to provide thematic continuity, helping users understand what they were looking at after a screen change.

Many of the examples in this Guide used fairly simple objects and groups, so that the relationships between the figures, the structure diagrams and the transition path diagrams were not too complicated. The MATIS example illustrated how more complicated window-based display designs can be analysed using the same techniques.

Hints for structuring displays

1 Consider what objects you want to appear on the display.
Does the presentation on the screen actually form these objects?
Are the individual objects recognisable as propositionally nameable entities?

2 Do the objects on the display form groups, or a hierarchy of groups?
Do you have an array of icons, or a display containing smaller windows which in turn contain further structures?
Draw the **Structural Diagram**, and be careful to consider the structure of the different objects, and whether these contain smaller objects that the user can interact with.

3 If the user has to search for objects in the display, what mental image (object representation) will they form: what attributes of an object can they use to discriminate it from its neighbours?
Do these attributes help the target pop out from the background?
Remember that a user might use propositional knowledge for different purposes, and so the structure of the object representation might vary.

4 When the structure diagram is drawn, think about the visual transitions users will make, and draw a **Transition Path Diagram**.
What sequences of tasks are users going to perform with the display? What will the transition paths look like?
Are there ways you could speed up the tasks by avoiding circuitous transitions between levels of structure, e.g. designing structures so that the important objects are grouped?

Sources of examples

The icon arrays shown in Figures 1, 2, 25, 27, and 29 are all taken from the Macintosh Finder, System 7.1.1.

The menu shown in Figure 20 is a (customised) version of one found in Microsoft Word 5.1a for Macintosh.

The representational and abstract icon sets shown in Figure 22 were first used by:
Arend, U., K-P. Muthig and J. Wandmacher (1987) 'Evidence for global feature superiority in menu selection by icons', *Behaviour and Information Technology, 6*, 411-426.

Other experiments carried out using these icons are reported in:
Green, A.J.K. and P.J. Barnard (1990) 'Icon Interfacing: The role of icon distinctiveness and fixed or variable screen location', in D. Diaper, D. Gilmore, G. Cockton and B. Shackel (Eds) *Proceedings of Interact '90*, Amsterdam: Elsevier Scientific Publishers B.V., pp 457-462

The analysis of the structure of the icons, and the graph shown in Figure 24, appeared in:
May, J., Barnard, P.J. and A. Blandford. (1993) 'Using Structural Descriptions of Interfaces to Automate the Modelling of User Cognition', *User Modelling and Adaptive User Interfaces, 3,* 27-64.

The 'cube' and 'empty box' objects shown in Figure 32 were first described by:
Enns, J.T. & Resnick, R.A. (1992) A model for the rapid interpetation of line drawings in early vision. In D. Brogan (Ed.), *Visuial Search II*, pp. 73-89. London: Taylor & Francis.

The hypertext database shown in Figure 39 was developed by researchers at the University of York, and is described in:
Myers, K.J. and N.V. Hammond (1991) 'Consolidated Report of workshop on scenario matrix analysis'. Esprit 3066 'Amodeus' Deliverable D9, Dept. of Psychology, Univ.of York, UK

The 'face icons' of Figure 49 were described in:
Chernoff, H. (1973) The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association, 68*, 361-368.

The 'stick men' icons of Figure 51 were developed by:
Pickett, R.M. & Grinstein, G.G. (1988) Iconographic displays for visualising multidimensional data. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics.* Beijing and Shenyang, PRC: IEEE.

The 'MATIS' system shown in Figure 56 is described in:
Nigay, L. & J. Coutaz, (1995). A generic platform for addressing the multimodal challenge. In *Proceedings of CHI'95*. ACM: New York (in press).

Index of Exercises

Acknowledgements

The authors of this guide would like to thank Anker Jørgensen and Annette Aboulafia of the Psychology Department of Copenhagen University, and their many HCI students for their help in the development and evaluation of earlier drafts.

This Guide was written as part of the AMODEUS project: ESPRIT basic research action 7040, supported by the Commission of the European Union.

Index of key points