2004

Time course of perceptual grouping in user interface displays

Melissa F. Schulz
University of South Florida

Follow this and additional works at: http://scholarcommons.usf.edu/etd
Part of the American Studies Commons

Scholar Commons Citation
Schulz, Melissa F., "Time course of perceptual grouping in user interface displays" (2004). Graduate Theses and Dissertations. http://scholarcommons.usf.edu/etd/1239

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Time Course of Perceptual Grouping in User Interface Displays

by

Melissa F. Schulz

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy Department of Psychology College of Arts and Sciences University of South Florida

Major Professor:  Doug Rohrer, Ph.D Emanuel Donchin, Ph.D. Paul McCright, Ph.D. Dewey Rundus, Ph.D. Toru Shimizu, Ph.D.

Date of Approval:
May 2, 2004

Keywords:  Screen Layout, Organization, Global Processing, Local Processing

© Copyright 2004, Melissa F. Schulz
Table of Contents

List of Tables iii
List of Figures iv
Abstract vi
Time Course of Perceptual Grouping In User Interface Displays 1
Organizing the complex visual stimulus 7
  Bottom-up processing 8
  Top-down processing 9
Assumptions 12
Gestalt grouping in user interface 16
Time limited task demands 24
  Meeting deadlines 24
  Information integration 26
  Unwanted information sorting 29
  Repetition of familiar tasks 31
  Size perception 35
  Perception of global and local information 36
  Perception of coarse and fine scene properties 38
Object completion 42
  Perception of lightness and transparent filters 45
  Grouping by proximity and alignment 46
  Grouping by chromatic color similarity 47
  Grouping by achromatic lightness 53
  Grouping by shape similarity 55
Current line of inquiry 58
General Method 70
Participants 70
Design 70
Stimuli 71
  Ready signal 71
  Primes 71
  Targets 72
  Practice stimuli 73
Mask 74
Experiment 1
Method 76
Results and discussion 78

Experiment 2A and 2B
Method 84
Results and discussion 87

Experiment 3
Method 94
Results and discussion 95

General Discussion 99
Contribution to engineering 103
Contribution to vision science 107
Future directions 108
Overall summary 110

References 112

About the Author
End Page
List of Tables

Table 1  Characteristics of RT distribution for Experiment 1  82
Table 2  Characteristics of RT distribution for Experiment 2A  90
Table 3  Characteristics of RT distribution for Experiment 2B  92
Table 4  Characteristics of RT distribution for Experiment 3  97
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Illustration of context effects that affect letter interpretation</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>Gestalt grouping principles shown in a single line of dots</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Gestalt principles of grouping shown by lines</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>Gestalt principles of grouping that have recently been discovered</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>Narrative and structured format experimental interface</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>Computer program assessment of narrative and structured format</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>Icons in visual search experiment</td>
<td>21</td>
</tr>
<tr>
<td>8</td>
<td>Database with ambiguous grouping</td>
<td>26</td>
</tr>
<tr>
<td>9</td>
<td>Flight status display with ambiguous grouping</td>
<td>28</td>
</tr>
<tr>
<td>10</td>
<td>Pop-up advertisement with ambiguous grouping</td>
<td>31</td>
</tr>
<tr>
<td>11</td>
<td>Military program with ambiguous grouping</td>
<td>32</td>
</tr>
<tr>
<td>12</td>
<td>Experimental stimuli with global and local letters</td>
<td>37</td>
</tr>
<tr>
<td>13</td>
<td>Ambiguous chromatic color grouping stimuli and results</td>
<td>48</td>
</tr>
<tr>
<td>14</td>
<td>Unambiguous color grouping targets</td>
<td>52</td>
</tr>
<tr>
<td>15</td>
<td>Results for ambiguous priming study</td>
<td>53</td>
</tr>
<tr>
<td>16</td>
<td>Ambiguous achromatic color grouping stimuli and results</td>
<td>54</td>
</tr>
<tr>
<td>17</td>
<td>Ambiguous shape grouping stimuli and results</td>
<td>56</td>
</tr>
<tr>
<td>18</td>
<td>Ambiguous website</td>
<td>59</td>
</tr>
</tbody>
</table>
Figure 19. Pilot stimulus 1 and results 61
Figure 20. Pilot stimulus 2 and results 63
Figure 21. Pilot stimulus 3 and results 64
Figure 22. Primes used in reported experiments 67
Figure 23. Targets used in reported experiments 73
Figure 24. Practice stimuli used in reported experiments 73
Figure 25. Experimental and control sequences for Experiment 1 78
Figure 26. Results for Experiment 1 80
Figure 27. Additional analysis for Experiment 1 data 81
Figure 28. Targets for Experiment 85
Figure 29. Experimental sequences for Experiment 2 86
Figure 30. Results for Experiment 2 88
Figure 31. Additional analysis of Experiment 2A data 89
Figure 32. Additional analysis of Experiment 2B data 91
Figure 33. Experimental sequence for Experiment 3 95
Figure 34. Results for Experiment 3 96
Figure 35. Improved database 102
Perceptual grouping describes the organization of small elements into larger objects. Research in user interface (UI) design has demonstrated effects of perceptual grouping on attention and navigation. However, grouping can be mediated by a variety of task factors. One such mediator is processing time. Recent discoveries in vision science suggest that elemental grouping can occur in more than one way, depending on how long elements are displayed. These findings have led to a new understanding of perceptual organization of elements in real-world spatial environments. However, these findings had not been explored within the context of UI environments. Time limits to UI are often set by task demands. Exposure time limits may affect perceptual grouping of elements in UI. Here I report a series of experiments that tested global and local pushbutton grouping by time in user interface displays. The research question was to determine whether global or local depictions of pushbutton groupings speed interaction with user interface. Global and local groupings were compared because prior researchers have discovered that global scene properties can be perceived before local scene properties. For this reason, it was hypothesized that global, as opposed to local, depictions of pushbutton groupings would speed human-interface interactions. Global grouping was defined as grouping by
relatively large shapes whereas local grouping was defined as grouping by shapes that were relatively small. The difference between global and local grouping was quantitative and defined by comparison. Participants saw pushbutton interface displays on a computer monitor for varied exposure durations and were asked to make decisions about the grouping of pushbuttons in these displays. Responses and reaction times were recorded. The results of the reported experiments suggest that global, as compared to local, groupings are more accessible across stimulus durations. They also suggest that global groupings can be utilized faster than local groupings in unlimited exposures. Taken together, the reported results further our understanding of global and local Gestalt grouping in user interface displays.
In Donald Norman’s “The Design of Everyday Things,” a coffeepot by Jacques Carelman illustrates the notion that some designs simply work better than others. At first glance, we may think that Carelman’s coffeepot looks normal; it has a body, a handle, a lid and a spout. But at second glance, we realize that the coffeepot has a great design flaw. Namely, the spout and the handle are on the same side. This coffeepot is a simple demonstration that even in instances in which an object has very few parts, the design and organization of these parts can make a dramatic difference on how useful the object can be.

Our world is filled with countless examples of how an object’s usefulness can be determined by the organization of its design. Modern computers are no exception. Computers with well-organized functions are easier to interact with and easier to use. But how can we define the optimal organization for modern computer functions? The answer to this question is not entirely clear. As a society, we have witnessed a rapid progress in computerized technology that has yielded a multitude of newfound capabilities. However, the realization of this progress has permitted us little time to think about how to best organize these capabilities, hastily compromising the usability of these machines.
To help users handle the increasing functional complexity in modern computers, a common design strategy has been to add even more complex functions to these machines, like artificial intelligence, to ‘supervise’ the chaos (Cooper, 1999). This method does not seem like a sound solution in the grand scheme of things, because it fixes functional complexity by adding functional complexity. For this reason, a new method in computer design is needed, one that will concentrate on the way humans see and interact with computers (Card, Moran & Newell, 1983; Cooper, 1999; Gardiner & Christie, 1987; Norman, 1992; Norman, 1993; Reeves & Nass, 2000).

A human-centered approach to computer design requires an understanding of the means of communication between humans and machines. Computers comprehend machine languages that are constructed of strings of ones and zeros (Deitel, Deitel, & Nieto, 1999). Humans comprehend spoken and/or written languages that are constructed of alphanumeric characters. The challenge that designers face is that computers can only directly understand machine languages and machine languages are much too complicated for most humans to use proficiently (Deitel & Deitel, 2001). To resolve this problem, modern computers are programmed so that humans do not interact with the computer by direct manipulation of the binary machine code. Rather, programmers have created an indirect method of communication which bridges the linguistic gap between humans and machines, user interfaces.

User interfaces are defined as the part of the computerized technology that humans interact with, including things like hyperlinks, pushbuttons and scrollbars. The design of computerized user interfaces involves determining what the computer display
will look like and how it will communicate with the computer user (Cooper, 1995). User interfaces are like translators that help humans and computers interact with one another. Of course, we are only as good as our translator and certain interface designs work better than others. For this reason, the design of a user interface must be planned. Since the 1960s, the evolution of user interface has continuously advanced toward a human centered approach to design. Our interactions with modern desktop environments have come a long way from interactions with archaic binary code.

The most primitive type of interface used batch coding, which was prevalent in the 1960s. This type of coding involved using binary machine code to execute functions. Mainstream users did not actually use binary codes. Instead, they would ‘interface’ with a computer operator and the operator would input the codes for them. For the mainstream user, human-computer interaction was mediated completely by a human third party. Clearly, having to go through a third party limited the users’ control and flexibility when working with a computer. A change in design was needed, one that allowed the mainstream user to more frequently interact with computerized machines.

To make it possible for mainstream users to interact with computers, command line coding was created. With command line coding, the mainstream computer user could type in a word to execute a function that meant something to humans, rather than typing in strings of ones and zeros. For example, the command ‘print’ was created. When users typed in the word ‘print’, the computer would print the name of the document that followed this command. The advantage of the command line user interface was that there was typically an intuitive mapping between the command that was to be typed by the user
and the function that the computer would perform. The command line language was much easier for humans to learn and use, because it was similar to their natural language. But how could programmers make this interaction even more seamless?

To enhance human-computer interaction even further, graphical user interface (GUI) began to take shape during the 1980s and 1990s. GUIs display icons that metaphorically represent computer functions. Instead of typing the word ‘print’ to achieve the print function, for example, users click an icon on the screen with a printer on it. GUIs made user interfaces even more usable and more accessible to many people, because there was a metaphorical mapping between functions and ideas. Further, users did not have to try to remember commands to interact with the GUI proficiently. Rather, users only needed to recognize functions displayed by icons that represented these functions.

Despite the progress of the past, there is still much to improve with user interface design. Each paradigm shift in user interface design, coupled with decreases in cost, has contributed to a dramatic increase in computer users (Robertson, 1998). So what will the next paradigm shift be for user interface design? How can user interface be further enhanced?

There are several key areas of focus when it comes to enhancing user interface design. Researchers and developers study how to improve the effectiveness, efficiency and safety of their interfaces (Preece, Rogers & Sharp, 2002). They examine how learnable their interfaces are and how easy it is for users to remember how to interact with these interfaces (Preece et al., 2002). Researchers and developers also focus on the
visibility of important interface elements and screen areas, the quality of feedback provided by the interface, and the consistency of the interface across different areas of the same program or suite (Norman, 1988). They seek to create a good match between the interface and the real world, to help users recognize and recover from errors and to design interfaces that prevent errors from happening (Nielsen, 2001). They also seek to make interfaces handicap accessible and friendly across cultures (Nielsen, 2000). Of course, researchers and developers are also interested in finding ways to enhance user experience with interfaces. They try to make interfaces more satisfying, entertaining, fun and aesthetically pleasing to use (Preece et al., 2002). As one can see, there are a variety of key areas of focus in modern user interface design. Progress in these areas will come together to further enhance human-computer interaction.

One research topic that is relevant to many of the previously highlighted areas of focus is perceptual grouping. Perceptual grouping describes the process by which image elements are seen as belonging together (Palmer, 1999). Perceptual grouping is relevant, for instance, because it can contribute to visibility, making important elements or areas in the interface more defined. It can also be used to decrease user errors by highlighting safe areas of the screen and de-emphasizing other unsafe areas. Grouping can increase ease of use and efficiency by creating more definition between elements on the screen, improving navigation. It can also enhance the aesthetics of the interface, thereby making the interface more pleasing to use. Therefore, an examination of research in perceptual grouping as it relates to user interface design may provide a relevant and useful method
of achieving a more seamless interaction between humans and machines. So how much progress has been made in this area of study?

Researchers are beginning to shed light on the effects of perceptual grouping on attention and navigation in interface environments (Addy, 2000; Bennet, Nagy & Flach, 1996; Card, 1982; Niemela & Saarien, 2000; Proctor & Proctor, 1996; Tullis 1981; Tullis, 1984; Tullis, 1986; Wickens & Carswell, 1996). However, recent discoveries in vision science suggest that perceptual grouping can be moderated by exposure time (Kurylo, 1997; Schulz, 2001; Schulz, 2002; Schulz, Peterson, Sanocki & Sellers, 2001; Schulz & Sanocki, 2003). These findings have not been considered within the context of user interface design, but are of great interest because exposure time to user interface is often limited by task demands. In light of recent findings and due to time limits set by task demands, it is essential to consider the effect of exposure time on perceptual grouping in user interface design.

This review serves to highlight recent discoveries in perceptual grouping and to describe how these findings can be applied to user interface design. First, a discussion about how humans organize the complex visual stimulus will be presented with specific emphasis on Gestalt grouping. Second, Gestalt grouping in user interface will be detailed. Third, time limits set by task demands in user interface will be described. Fourth, recent discoveries in perceptual organization will be highlighted. Finally, in light of this review, a new line of research will be discussed that has tested the potential application of recent discoveries in Gestalt grouping to user interface design. Specifically, a series of experiments that tested global and local pushbutton grouping by time in user interface
displays are reported. Global and local grouping are compared in the present study because researchers have found that global scene properties can be perceived before local scene properties across time (Navon, 1981, Sanocki, 1993). Theorists have defined global scene properties as being the largest size scale shapes of objects and perceptual scenes (Sanocki, 2001; Navon, 1977; Navon, 1981) whereas local scene properties have been defined as being more internal, interior and generally smaller scale shapes in perceptual scenes (Sanocki, 1993). In the present study, global grouping was defined as grouping as being relatively large in size scale when compared to local grouping. Here, the difference between global and local grouping was defined quantitatively.

*Organizing the complex visual stimulus*

Human vision involves much more than simply opening our apertures and visually perceiving the world. The 2-dimensional image that is cast onto the retina of the eye requires a certain amount of processing to generate the rich 3-dimensional world that we perceive (e.g., Goldstein, 1999). This processing is essential because human visual perception is not as automatic and effortless as it may at first seem. The retinal receptors of the human eye detect local pieces of spatial scenes. These individual local pieces, that together compose the 2-dimensional image stimulus, are ambiguous when considered independently due to scene attributes like shadows, real edges, and colors. In a more general sense, image formation can be thought of as a many-to-one mapping (e.g., Nalwa, 1993). Given any retinal image feature, there are a variety of distal spatial scenes that could have produced it. Yet despite these factors, humans rapidly generate global
interpretations of perceptual scenes that are constructed of the ambiguous retinal pieces. So how does the ambiguous local information that our retinal receptors detect become the global scenes that we perceive?

To construct global scenes, humans must rely on visual processing, which integrates the visual information that the retinal receptors detect with prior knowledge about the spatial environment and a variety of assumptions. In what follows, these integrated components of the visual process are examined individually. Bottom-up, top-down and assumption based visual processing will be discussed. Specific emphasis will be placed on Gestalt grouping assumptions due to their central role in the proposed research.

**Bottom-up processing.** Bottom-up image processing begins when a patterned array of light from the environment strikes a complex network of cells that covers the inside back of the eye. This network of cells is called the retina and includes receptors that fire electrical signals in response to light (Goldstein, 1999). Retinal receptor cells fire in various configurations, depending on the configuration of the patterned array of light striking the eye. When the retinal receptor cells generate electrical signals in response to an image, the visual system receives its first piece of information.

Initial pieces of visual information are sent from retinal receptors to ganglion cells that form optic nerves. The optic nerves transmit the information to the lateral geniculate nucleus (LGN) of the thalamus. The LGN then sends the electrical message to the cortical receiving area of vision, in the occipital lobe of the brain for further processing (Goldstein, 1999).
Electrical activity generated by the retinal image is initially registered in the cortical receiving area as edges, lines, and blobs. Collectively, these representations have been described as raw primal sketches (Marr, 1982). Raw primal sketches represent information about the impoverished 2-dimensional structure of the retinal image, rather than rich information about the physical objects in the external world that produced that image (Palmer, 1999). For example, a red apple will be represented in the cortical receiving area as being purple if it is viewed by an observer in dim lighting conditions. The apple appears to be purple because the color of the apple actually detected by the retinal receptors is purple when viewed in these conditions. However, the color of the apple detected in these conditions is not that which we know a red apple to be.

To become the red apple, the electrical activity representing bottom-up visual information about this object must be sent from the primary receiving area of the visual cortex as output to many deeper visual areas (Palmer, 1999). When the electrical activity reaches deeper levels of the visual cortex, it is combined with stored knowledge that helps humans interpret the ambiguous bottom up visual information, represented as the electrical raw image information. The application of this stored knowledge to the bottom-up stimulus information, toward the development of final percept, is known as top-down processing.

*Top-down processing.* Bottom-up processing generates only one type of information that humans rely on to interpret spatial scenes. By itself, it is simply not adequate because it does not contain enough detail and can often be ambiguous as to how it should be interpreted. Humans must combine the information that they have attained
from bottom-up processing with prior knowledge about the visual world (e.g., Biederman, 1987; Peterson, 1994; Shepard, 1983). The modification of bottom-up information by application of prior knowledge is considered to be a top-down process.

One type of top-down processing is the use of context for image identification. When humans are familiar with the context of an image, they are often able to recognize this image at a faster rate (e.g., Palmer, 1975). For example, a neighbor outside her house might be recognized more quickly than if she were in a shopping mall. When seeing a neighbor in her house, humans apply what we know about who lives in that house to determine who is there. It is very easy to make this determination, even if our neighbor appears far away, because the object person is in a context where we often see her. However, when a neighbor is spotted in an alternative location, she may be more difficult to discern without the familiar context. A second example of the influence of context is the word superiority effect (Reicher, 1969; Wheeler, 1970). A target letter is identified more rapidly when primed within a word than when primed independently. In this case, the context word prime allows humans to more quickly interpret object letters when it is present. Without the primed word context, letters are less quickly identified.

The use of familiar context is not the only way that knowledge can influence what humans visually perceive. Another factor to consider is the schemata that humans have about particular spatial scenes. A schema is a pattern of knowledge that describes what is typical in a particular situation (Reisberg, 2001). We interpret objects in scenes according to the schema that we have about those scenes. For example, an office schema would specify that a desk and a chair would likely be present whereas a refrigerator and an oven
would not be. It would be easier to identify a desk in an office than to identify a desk in a kitchen. This is because the desk seems to be a likely object that we would find in an office setting, according to the schema that humans have about offices.

Sometimes the influence of top-down processing can be so strong that we can actually misinterpret bottom-up visual information. For example, in Figure 1, the middle letter of each word is interpreted differently based on its context. The middle letter in first word is perceived as an ‘H’ because that is the letter we expect in the word most visually similar to it, ‘THE.’ However, the middle letter in the second word is perceived as an ‘A’ because that is the letter we expect in the word most visually similar to it, ‘CAT.’ Both letters are actually the same size and shape and are neither an ‘A’ nor an ‘H.’ In this case, humans rely so much on the word context that the missing information is added by the brain. Similarly, humans may see objects in a scene that are not really there because these objects are part of the general schema for that scene (e.g., Intraub, Bender & Mangels, 1992; Intraub & Richardson, 1989). For example, we may interpret a large rectangular object in an office scene as a desk, even if it is really an oven, because we expect a desk to be present in an office.

Figure 1. Illustration of context effects that affect letter interpretation

Note. Middle letters in each word are identical in shape and size. However, humans perceive an ‘H’ in first word and an ‘A’ in second word due to expectations.
As one can see, top-down processing can be used to resolve the ambiguity that bottom up information contains. It can help humans interpret stimuli at a faster rate, without having to process all of the details. Without top-down processing, humans would spend a tremendous amount of energy processing the visual stimulus. Top-down processing makes vision more efficient in this way. However, top-down processing is not the only thing that makes vision more efficient. Another set of factors that facilitate the visual processing of spatial scenes are the innate assumptions that humans have about the visual world.

Assumptions. In addition to using bottom-up and top-down processing to organize information in spatial scenes, humans also rely on a variety of assumptions to interpret information. These innate assumptions appear to be similar across environments and cultures, and are quickly realized with human experience of the world (Hergenhahn & Olson, 1995). One subset of assumptions made in visual processing is based on the Gestalt principles of grouping (Koffka, 1922; Koffka, 1935; Köhler, 1924; Wertheimer, 1950). The Gestalt grouping principles aid humans in the perception of objects in the retinal image input that they detect. These principles are central to the proposed line of research. For this reason, an overview of the Gestalt grouping principles will be given with emphasis on how these principles have guided our understanding of human vision in real world spatial environments.

In the early 1920s, a group of researchers began to discover the underlying principles that humans use to organize their perception of the world (Koffka, 1922; Koffka, 1935; Köhler, 1924; Wertheimer, 1950). These researchers showed that elements
could be grouped by the principles of proximity, similarity and common fate (see Figure 2). In panel A of Figure 2, for example, there is no grouping pattern because the dots are identical to one another and evenly spaced. For this reason, a single row of independent black dots is perceived. In panel B, however, the dots have been paired by location so that there are two groups of dots, which demonstrates the grouping principle of proximity. In panel C, the dots are paired by similar colors, which demonstrates grouping by the principle of similarity. In panel D, the two moving dots are grouped separately from a second group of two static dots, thereby demonstrating grouping by common fate.

Figure 2. Gestalt grouping principles shown in a single line of dots

Note. (a) Single row of evenly spaced dots. (b) Grouping by proximity. (c) Grouping by similarity. (d) Grouping by common fate.

In addition to the previously mentioned principles, other principles guiding the grouping of linear elements were discovered. In Figure 3A, two pairs of lines are mirror reflections of one another and therefore appear to go together, therefore demonstrating the principle of symmetry. In panel B, two sets of parallel lines form two groups of two lines, demonstrating grouping by the principle of parallelism. In panel C, the smooth continuous shape appears to be constructed of two lines, rather than four, as consistent
with the principle of good continuation (continuity). In panel D, two pairs of lines appear
to form two boxes that almost close, demonstrating grouping by closure.

Figure 3. Gestalt principles of grouping shown by lines

A. 
B. 
C. 
D. 

Note. (a) Grouping by symmetry. (b) Grouping by parallelism. (c) Grouping by good continuation. (d) Grouping by closure.

Contemporary researchers continue to discover new principles of grouping. The
principle of synchrony describes the idea that events that occur at the same time will be
grouped together (Bregman, 1978; Palmer & Levitin, in preparation). Due to its dynamic
properties, the principle of synchrony is not easily depicted in a static figure and therefore
not depicted in this paper. The principle of connectedness implies that elements that are
connected in some way tend to be perceived as belonging together (Palmer & Rock,
1994a; Rock, Linnet et al., 1992). In Figure 4A for example, the first and second pair of
dots are each connected by a black bar. Due to the presence of the connecting bars, the
dots appear to form two separate groups. The principle of common region describes the
idea that elements that are enclosed in the same region will be grouped together (Palmer,
1992). In Figure 4B, for example, there are two boxes, each surrounding a pair of dots.
Due to the presence of the boxes, two groups of two dots are perceived, rather than four
independent dots.
Figure 4. Gestalt principles of grouping that have recently been discovered

A. [Diagram]

B. [Diagram]

Note. (a) Grouping by connectedness. (b) Grouping by common region.

The Gestalt principles of grouping are presumed to be the underlying assumptions on which perceptual organization is based (Wertheimer, 1950). For this reason, the relationship between these principles and attention and navigation has been carefully considered with respect to human perception of real-world spatial environments (Beck, 1967; Ben-Av, Sagi & Braun, 1992; Duncan & Humphreys, 1989; Johnston, Schwarting, & Hawley, 1996; Moore & Egeth, 1997; Olson & Atteave, 1970; Pomerantz & Garner, 1973; Pomerantz & Schwitzberg 1975; Treisman, Sykes & Gelade, 1977; Treisman & Gelade, 1980; Treisman, 1982; Treisman, 1985; Yantis, 1992). The Gestalt principles of grouping have also had many diverse applications in a variety of domains including audition (e.g., Bregman, 1990), computer vision (e.g., McCafferty, 1990), and art (e.g., Arnheim, 1974). Given the diverse applicability of the Gestalt principles, it is reasonable to question the extent to which these principles could be applied to enhance user interface design. Does the applicability of the Gestalt principles of grouping extend well to the user interface domain?

Many perceptual phenomena that are observed in human-real world interactions have been observed in human-computer interactions as well (Nass & Moon, 2000; Reeves & Nass, 1996; Reeves & Nass, 2000). Human perception of size and distance
(Reeves, Lang, Kim, & Tatar, 1999), motion (Reeves, Thorson, Rothschild, McDonald, Hirsch & Goldstein, 1985), novelty (Geiger & Reeves, 1993; Reeves et al., 1985), scene change (Geiger & Reeves, 1993), and faces (Nass & Gong, 1999; Nass, Isbister & Lee, 2000; Nass, & Lee, 2000; Nass, Moon & Green, 1997) have been shown to occur in computerized interface environments in much the same way that they do in the real world. Based on these findings, it seems logical to think that the Gestalt grouping principles could determine the grouping of elements presented in computerized environments in much the same way that these principles determine human perception of real world elements. Researchers in user interface design have begun to apply basic Gestalt grouping principles to computerized displays. But how far have they taken this application? Have they taken it far enough?

**Gestalt grouping in user interface**

Designers are beginning to recognize the importance of applying the Gestalt principles of grouping to user interface (Bailey, 1982; Bellcore, 1995; Card, 1982; Danchak, 1976; Galitz, 1985; Holden, Adolf & Woolford, 1997; Jones & Okey, 1997; Moore & Fitz, 1993; Stewart, 1976; Streveler & Wasserman, 1984; Tullis, 1983; Tullis, 1988; Williges & Williges, 1981). Several researchers have tested the effect of basic Gestalt grouping principles in interface design (Card, 1982; Niemela & Saarien, 2000; Tullus 1981; Tullis, 1984; Tullus, 1986). This research has paved the way to an understanding of how to efficiently group icons, buttons, pictures, text and other elements within user interface displays.
One of the leaders in grouping and user interface research is Thomas Tullis.

Tullis (1981) began by evaluating several types of user interface displays found in the telecommunication industry. These displays were used to help operators make decisions about what was wrong with particular telephone lines. Tullis presented displays in several formats and tested the usability of each display type. Two of these interface display types included narrative format, which contained unstructured text (see Figure 5A), and structured format, which used grouping principles like common region and proximity to organize the text into chunks (see Figure 5B). While Figure 5A appears to be moderately grouped by proximity, it is far less organized than Figure 5B. Participants were presented with both narrative and structured format display types and were asked to answer a series of questions about each. In addition, participants were asked to rate the overall quality of each display format on a scale of one to seven.

Figure 5. Narrative and structured format experimental interface

<table>
<thead>
<tr>
<th>A</th>
<th></th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TEST RESULTS</strong></td>
<td><strong>SUMMARY</strong></td>
<td><strong>GROUND</strong></td>
</tr>
<tr>
<td>GROUND, FAULT T.-Q</td>
<td>3 TERMINAL DC RESISTANCE</td>
<td>3 TERMINAL DC VOLTAGE</td>
</tr>
<tr>
<td>* 3500 K OHM T.-R</td>
<td>* 24 K OHM T.-R</td>
<td>* 0.05 VOLTS T.-Q</td>
</tr>
<tr>
<td>* 35 K OHM T.-R</td>
<td>* 35 K OHM T.-R</td>
<td>* 0.05 VOLTS T.-Q</td>
</tr>
<tr>
<td>3 TERMINAL AC SIGNATURE</td>
<td>* 350 K OHM T.-R</td>
<td>0.05 VOLTS T.-Q</td>
</tr>
<tr>
<td>* 3500 K OHM T.-R</td>
<td>* 3500 K OHM T.-R</td>
<td>* 0.05 VOLTS T.-Q</td>
</tr>
<tr>
<td>LOW RESISTANCE</td>
<td>9 K T.R</td>
<td>14 K T.G</td>
</tr>
<tr>
<td>LONGITUDINAL BALANCE POOR</td>
<td>0 V T.G</td>
<td>0 V T.G</td>
</tr>
<tr>
<td>BALANCE</td>
<td>CENTRAL OFFICE</td>
<td>62 K RES.</td>
</tr>
<tr>
<td>VALID LINE C.T. CONFIGURATION</td>
<td>39.08</td>
<td>VALID LINE C.T.</td>
</tr>
<tr>
<td>CAN DRAW AND BREAK DIAL TONE</td>
<td><strong>END</strong></td>
<td>DIAL TONE OK</td>
</tr>
</tbody>
</table>

Note. Stimuli presented by Tullis (1981) (a) Narrative stimulus contains complete unstructured text. (b) Structured stimulus contains text that is grouped by closure and proximity.
Participants were asked to complete two experimental sessions. Although no differences in accuracy were observed, participants’ second session latencies were significantly briefer for the structured display when compared with the narrative display. In addition, across both sessions, participants assessed the quality of the structured displays to be significantly better than that of the narrative displays. These results suggest that participants are faster at making decisions about text that is logically grouped, when compared to unstructured text. The results also suggest that participants find structured displays to be of a superior quality when compared to unstructured displays.

When interpreting these results, it is important to note that the structured format display type used slightly fewer text characters than the narrative format, thus making the search set smaller. A smaller search set should produce faster visual search and this could explain the participants’ speeded responses to questions relating to the structured display. To understand grouping in interface displays, further research would be needed to tease these factors apart.

One of the major setbacks with the Tullis (1981) research was that it was not objective in quantifying grouping. Had grouping been quantified, the degree to which grouping existed in both experimental display types could have been better contrasted. To speak to this issue, Tullis (1984) developed a computer program to quantify six interface characteristics within his stimuli. Two such characteristics included “number of groups” and “average size of groups.” With respect to Figure 5A for example, Figure 6A shows the number of groups defined by the computer program for this stimulus. Groups are
defined by similar numbers. For example, all of the ones in Figure 6A form a group. As a second example, with respect to Figure 5B, Figure 6B shows the number of groups defined by the computer program for this stimulus. In comparing Figure 6A and Figure 6B, the Tullis computer program has assessed Figure 6A as having three groups whereas Figure 6B is assessed as having 13.

Figure 6. Computer program assessment of narrative and structured format

Note. Stimuli presented by Tullis. Groups are represented by identical numeric values (1984) (a) Three groups are represented within this stimulus. (b) Thirteen groups are represented within this stimulus.

Using this program to assess grouping in user interface, Tullis (1984) performed a usability study on displays with different numbers of groups. Participants were presented with stimuli that contained a range of computer-defined groups. Participants were asked to answer questions about each stimulus. Researchers recorded the amount of time it took for participants to locate the correct answers to these questions. The results suggested that search time increases with the number of groups and the average size of groups. This research is consistent with the idea that grouping in user interface displays can facilitate search time for finding answers to questions that are based on displays.
In an extension of his 1984 study, Tullis (1996) set out to determine the specific relationship between the number of groups and the average size (visual angle subtended) of groups in user interface displays. Specifically, Tullis was interested in the group number and visual angle range that optimized search speed. Tullis found that search time is a function of the number of groups when groups contain visual angles measured at less than 5 degrees. That is, for groups subtending a 5-degree visual angle, search time increased with the number of groups. However, when the visual angle exceeds 5 degrees, search time increased proportionally with visual angle.

Research by Tullis has assessed grouping interpretations and how fast users can answer questions about grouped and ungrouped interface displays. However, other lines of research have shown that grouping can affect users’ tasks with interface in more indirect ways as well. As one example, Niemela and Saarinen (2000) used a visual search task to study the effect of icon grouping on scanning speed. These researchers hypothesized that the spatial grouping of icons of the same application type would increase scanning speed.

To test this hypothesis, participants were presented with an icon-based user interface, for an unlimited exposure duration, and asked to search for a target icon located among four to sixteen distracter icons. In one condition, icons representing files that were saved in the same application type were presented in proximal groups (see Figure 7A). In a second condition, however, icons representing files saved in the same application type were randomly displayed (see Figure 7B). Across 4, 8, and 16 item set sizes, Niemela
and Saasrinen (2000) observed that participants were able to locate files from particular applications faster when icons representing these applications were spatially grouped (778ms) rather than randomly arranged (1267ms). In other words, the spatial grouping of icons of the same application type significantly increased scanning speed for all search set sizes.

Figure 7. Icons in visual search experiment

A

B


Niemela and Saarinen (2000) have provided information about how the use of independent Gestalt grouping principles like proximity and similarity can enhance user interface design. But what about interactions between Gestalt grouping principles? Do Gestalt grouping interactions affect display design? To speak to this question, additional research that focuses on how Gestalt principles are perceived together in complex screen displays has also been a topic of interest (Addy, 2000; Bennet, Nagy & Flach, 1996; Proctor & Proctor, 1996; Wickens & Carswell, 1996). This line of inquiry is important
because Gestalt grouping principles are often not used redundantly in interface displays. In many cases, these principles are mistakenly used in opposition, giving rise to ambiguities. One such example of grouping interaction research is a study by David Addy (2000) who compared the relative strengths of similarity, proximity and common region grouping in display design.

In Addy’s experiment (2000), each stimulus contained a single row of dots that could be organized by one of two different grouping principles. Participants were shown these stimuli on a computer screen for an unlimited amount of time and asked to report how they thought the dots should be grouped in each display. Across experiments, participants tended to group by color similarity, followed by common region, then followed by proximity. Addy (2000) had shown that, in certain conditions, some grouping principles could be stronger than others on computerized displays. Addy’s research exemplifies the importance of understanding the interactions between various Gestalt grouping principles in user interface design.

Born out of grouping and interface design research, theorists have generated ideas about how the Gestalt principles of grouping can most effectively be applied to interface displays. A popular modern approach has been the Proximity Compatibility Principle (PCP), which combines information-processing models with Gestalt principles of grouping to create guidelines for how information should be displayed (Barnett & Wickens, 1988; Carswell & Wickens, 1987; Wickens & Andre, 1990; Wickens & Carswell, 1995). The PCP expands on the engineering principle of functional grouping, in which functionally related instruments should be designed to be close in physical
proximity (Bailey, 1989; Bonney & Williams, 1977). Specifically, the PCP provides guidelines for the relationship between two measures: processing proximity and perceptual proximity. Processing proximity describes the extent to which multiple information sources are used for the same task (Wickens & Carswell, 1995). If these information sources are to be integrated for the task, then processing proximity is considered to be high (Wickens & Carswell, 1995). On the other hand, perceptual proximity describes the extent to which multiple information sources are perceptually similar. If these information sources are depicted proximally or are similar in color, physical dimension or coding, these sources are thought to have close perceptual proximity. The PCP suggests that if information sources have high processing proximity, then designers should implement high perceptual proximity of these sources in interface displays. Likewise, if independent processing is required of multiple information sources, then low perceptual proximity is advised for those sources in display design. In short, the PCP suggests that the level of display proximity should match the level of task proximity (Bennett, Nagy, & Flach, 1997).

Experimentation and principles about Gestalt grouping and interface design have provided researchers with important information about how to visually enhance the configuration of elements in displays. However, many of these experiments and principles have been conducted and developed under the assumption of unlimited time exposure to interface displays. Do unlimited exposure conditions necessarily match those that the typical user faces in completing a computerized task? The answer to this question is likely no. A typical users’ task may be speeded for a variety of reasons, causing their
perception of the interface display to be something other than the perception that
designers and experimenters analyze and implement. Recent research in vision science
suggests that the nature of grouping may be different in limited exposure time conditions
than it is in unlimited exposure time conditions (Kurylo, 1997; Schulz, 2001; Schulz,
2002; Schulz et al., 2001; Schulz & Sanocki, 2003). Based on these findings, it is critical
to consider the time limits set on tasks that users face when perceiving interface displays.

*Time limited task demands*

The relationship between Gestalt grouping and UI design has been carefully
considered in terms of attention and navigation (Bailey, 1982; Bellcore, 1995; Card,
1982; Danchak, 1976; Galitz, 1985; Holden, Adolf & Woolford, 1997; Jones & Okey,
1997; Moore & Fitz, 1993; Stewart, 1976; Streveler & Wasserman, 1984; Tullis, 1983;
Tullis, 1988; Williges & Williges, 1981). In these cases, Gestalt grouping has been
treated as a static property. However, grouping is not always static. There are several
moderating factors that can change our perception of grouping, making grouping much
more of a dynamic process. One such moderator is exposure time. User interface duration
can be limited by a variety of task factors. In what follows, a sample of these task factors
are briefly identified and evaluated.

*Meeting deadlines.* Computer users are rewarded for completing tasks quickly.
To accomplish this, they must navigate quickly through the interfaces that are displayed.
Consider the employee who is late for a presentation to which he had planned to bring a
picture from a database. The employee must quickly navigate through the database to
find the target ‘Picture’ in the ‘Presentation’ branch of this data structure. Designers of such databases know that well organized database elements will speed navigation and visual search for targets. For this reason, they group database elements in a way to optimize speeded visual search and navigation. But will the employee who rushes to meet a deadline perceive the same elemental grouping in limited interface exposures that the designer perceived when creating the database in unlimited interface exposures? The answer to this question is unclear. It is possible that the elemental grouping that the designer has implemented and perceived in unlimited exposure conditions may not be perceived by the user in speeded conditions. Further, it is possible that there exists an alternative elemental grouping interpretation perceived in speeded file search that works against the user by slowing search.

Figure 8 depicts a possible database structure that the employee in our example could be searching. Recall that the employee is looking for a target ‘Picture’ in the ‘Presentation’ branch of the database. Upon examining the structure of this database, it may appear as though finding the target ‘Picture’ would be a simple task. The elements within each branch of the database are grouped by the principle of connectedness. That is, each file is adjoined to its database by a line segment. If the employee perceives the elements as grouped by connectedness, he will likely reach his target ‘Picture’ very quickly. But what are the alternative grouping interpretations for the elements within this database? While the target file appears to group by connectedness with the files in the ‘Presentation’ branch, it also appears to group with by common region with the files in the ‘Personal’ branch of the database. Printing a target picture from the ‘Personal’ branch
of the directory to present at a meeting might prove embarrassing. So which interpretation would be perceived at short interface exposures? If the employee perceived the connectedness grouping interpretation, there would be no problem. However, if the employee perceived no grouping or the common region grouping of the file with the wrong branch of the database, his search would be slowed.

Figure 8. Database with ambiguous grouping

![Database with ambiguous grouping diagram]

*Note.* The ‘Picture’ element groups with the ‘Presentation’ branch of the database by connectedness and with the ‘Personal’ branch by proximity.

The above example serves to demonstrate how deadlines can limit exposure time to the elements displayed in user interface. It also serves to show how these limits could potentially moderate our perceived interpretations of the grouping of these elements. Deadlines are not the only task factor that can limit exposure time, however. Time may be limited by a variety of other types of task factors. One such factor is the process of information integration.

*Information integration.* Often times, users are asked to complete tasks that require them to integrate several pieces of information in order to make a decision. These pieces of information may be presented in several areas of one interface or on several different interfaces. In order to gather the information for a decision, users must saccade from one area or interface to the next and gather information in between fixations.
Humans integrate many types of information during saccadic eye movement (Irwin, Carlson-Radvansky & Andrews, 1995) including multiple object features like color and orientation (Carlson, Covell & Warapius, 2001), surface and edge properties of items (Gilchrist, Findlay & Heywood 1999) and lexical information (Inhoff & Tousman, 1990). In many tasks, the fixations can be as short as 200 milliseconds (Trappenberg & Klein, 1999). This is a very brief amount of time to perceive the organization of information that is displayed.

Consider the example of an airline pilot who must quickly scan information from a series of monitors to determine altitude and speed information before making a decision about landing. As an illustration, Figure 9 depicts a hypothetical interface arrangement for a pilot in which two interfaces are adjacent. In this example, there are two interfaces presented side by side. The interface on the left shows information about altitude while that on the right shows speed information. The pilot’s task is to locate the ‘Current’ measurement of altitude and speed to make a decision about landing. To make her decision, the pilot must quickly saccade across both interfaces to gather the relevant information. In unlimited exposure product development conditions, the designers of such interfaces use perceptual grouping to organize the information. These groupings are designed to increase the efficiency of information search. But are these groupings optimal for the time-limited eye saccades that take place during the pilot’s speeded information search?
Figure 9. Flight status display with ambiguous grouping

![Altitude and Speed Display](image.png)

*Note.* In each interface, text groups by the principle of color similarity within the ‘Status’ and ‘Measure’ columns. However, text also groups by alignment within the ‘Current’ and ‘Recommended’ rows.

The designer in our example has grouped the information in each display by the principles of alignment and color similarity (see Figure 9). Grouping by color similarity is used to show difference between the ‘Status’ and ‘Measure’ categories of information. Grouping by alignment is used to show the difference between the ‘Current’ and ‘Recommended’ categories. Recall that the pilot’s task is to locate the ‘Current’ measurement of altitude and speed before making a decision about landing. If the pilot groups the information by alignment, she will likely find the ‘Current’ status of both altitude and speed very quickly. However, the nature of the pilots’ task requires her to saccade across the displays very rapidly. In these conditions, is the designers’ intended grouping of information perceived? If so, is one grouping interpretation more salient than the other? If the alignment grouping were most salient in limited exposure conditions, then the pilot would locate the relevant information very efficiently. However, if color grouping where the most salient in limited exposure conditions, the scene would appear to be grouped by ‘Status’ and ‘Measure’ categories. In this case, the grouping of information would not help the user discriminate between ‘Current’ and ‘Recommended’
measures. As a result, the pilot would have to slow search for the relevant information. Clearly, this delay may have dire consequences. This example illustrates that the quick eye saccades, which are necessary for information integration, may limit exposure time to user interface. This example also shows how limits on exposure time could potentially affect the way users perceive the grouping of information in such situations.

In the previous example, the users try to attend to presented information. But what about situations in which the user’s goal is to not attend to the presented information? Are the effects of exposure time limits on perceptual grouping relevant in these settings?

*Unwanted information sorting.* Users are often presented with information that they do not wish to view. A good example of this type of information is that which is presented on Internet pop-up windows. How many milliseconds does it take to close a pop-up window? How long do the users actually see the information that is presented? Often times, the users only see the information presented in pop-up windows for a very brief time period because they are almost always more interested in searching for the close button than they are in scanning the advertisement. Designers know that the more meaningfully they group the information within pop-up windows, the more likely the users who are presented with such windows will be to correctly interpret this information. But the designers of pop-up windows may have weeks or even months to develop their displays. Are the information groupings that designers perceive in unlimited window development conditions the same as those that users perceive when rapidly attempting to close these windows?
Consider the case of an Internet surfer who is presented with a pop-up advertisement, like the one depicted in Figure 10. The advertisement lists four qualitative descriptions about a gym, and each is presented in the form of underlined textual links. In addition to these links, the advertisement presents a vital fifth piece of information, the ‘Join Today’ link. This link will guide users to a location where they can register as new members of the advertised gym. The ‘Join Today’ link may be the most important piece of information in the display and the one that should be presented the most clearly for the user. To accomplish this, the designer organized this advertisement so that the ‘Join Today’ link is grouped against the less relevant background information by the Gestalt principle of closure, within the frame of a push button. If the user should use closure to group the information in this advertisement, it is likely that she will identify the “Join Today” link right away. It should be noted, however, there are other possible grouping alternatives within this advertisement. While the “Join Today” link groups well by closure within the frame of the push button, it also groups by similarity with the underlined textual links that surround it. If the user were to group the “Join Today” link by textual similarity, there would be many more pieces of information to sort though before she could locate the essential ‘Join Today’ link. In our unlimited exposure to this example of an advertisement, it is easy to perceive both grouping interpretations mentioned. But which interpretation, if any, would be perceived in limited exposure conditions? And how does this interpretation affect the likelihood that the user will identify and click on the ‘Join Today’ link?
From the pop-up advertisement example, it becomes clear that unwanted information sorting can limit exposure time to user interfaces. What is not clear is how the information in these interfaces is perceived and grouped within these time limits. The pop-up advertisement example demonstrates a situation in which the user is viewing the information within an interface for the first time. But what if users have prior experience with the information on a display? In these cases, are concerns about time limitations relevant?

*Repetition of familiar tasks.* When users observe an interface that they have seen previously and are traveling back through it, they may not require as much time as those who are using the interface for the first time. It is likely that experienced users travel through familiar interfaces as quickly as their memory will allow them. Do these types of users perceive the same elemental grouping as first time users?
Consider the example of a military program that allows soldiers to preview a variety of different maps, as depicted in Figure 11. In this example, the pushbuttons near the top of the interface allow the soldiers to select from five different maps. In contrast, the pushbuttons on the right represent the Up, Down, Right, and Left functions that the soldiers can use to navigate through these maps. The designer of this military program used principle of proximity to separate the pushbuttons into two groups that each have a specific type of function: display map or navigate. In addition to the intended proximal grouping, both sets of pushbuttons also group against the gray background by color similarity and closure.

Figure 11. Military program with ambiguous grouping

Note. Pushbuttons group by the principle of proximity into two separate clusters. Pushbuttons also group by color similarity and closure against the gray background.

Soldiers who have prior experience with this military program will likely navigate through this interface very quickly. With a series of clicks and screen flashes, these users will rapidly locate their destination. So what type of pushbutton grouping, if any, do the
soldiers see in these speeded conditions? The answer to this question is not entirely clear. It is possible that the soldiers will perceive the two intended proximal pushbutton clusters. In this case, visual search for a target pushbutton will be speeded because the set size to search is reduced by half. The user simply picks the desired proximal group to search within and begins to scan that group. But what if the proximal grouping was not apparent in limited exposure conditions. What if, the pushbuttons were instead perceived only as being grouped against the gray background by color similarity and closure? In this case, the set size for visual search would not be reduced into two functionally similar groups. Rather, the soldiers would have to search through all of the pushbuttons to find the target function. It should be noted that, in speeded conditions, soldiers with prior experience could rely on other cues, like positional constancy, to locate the target pushbuttons. However, this example serves to demonstrate how the physical grouping may not provide any useful information, despite supplementary cues that might aid the user.

From all of the previous examples, it becomes clear that there are a variety of task factors that can limit exposure time to user interface. Limits on exposure time require users to process the grouping of the elements in interface displays very rapidly. In these time-limited conditions, is the organization that users perceive the same as in unlimited conditions? Recent discoveries in vision science would suggest that this is not the case (Gulick & Stake, 1957; Kurylo, 1997; Moore & Brown, 2001; Navon, 1981; Oliva & Schyns, 1997; Rauschenberger & Yantis, 2001; Sanocki, 1993; Schulz, 2002; Schulz et al., 2001; Schulz & Sanocki, 2003; Schyns & Oliva, 1994). Rather, recent discoveries in
vision science suggest that perceptual organization, and specifically perceptual grouping, can vary with exposure duration.

*Time course of perceptual organization*

Does perceptual organization occur immediately in vision or does it take time to manifest? There has been some evidence suggesting that scenes are organized so quickly that humans can access only organized structures, rather than the primitive elements that compose these structures (Rensink & Enns, 1995). However, several lines of recent research in vision science have shown that perceptual organization of elements in spatial displays can depend largely on how long these displays are presented (Gulick & Stake, 1957; Kurylo, 1997; Moore & Brown, 2001; Navon, 1981; Oliva & Schyns, 1997; Rauschenberger & Yantis, 2001; Sanocki, 1993; Schulz, 2002; Schulz et al., 2001; Schulz & Sanocki, 2003; Schyns & Oliva, 1994). This research has shown that perceptual organization may operate along a time course in which certain grouping interpretations are perceived under shorter stimulus durations but then changed by alternative grouping interpretations that are perceived at longer durations.

In order to understand how humans perceptually organize spatial and computerized displays, the perceptual process must be disrupted at different times. Researchers who have disrupted the visual process by limiting stimulus exposure time have found effects of stimulus duration on the perception of size (Gulick & Stake, 1957), low versus high frequency information (Schyns & Oliva, 1994, Oliva & Schyns, 1997), global versus local configuration (Navon, 1981; Sanocki, 1993), object completion (Rauschenberger & Yantis, 2001), lightness and transparent filters (Moore & Brown,
grouping by proximity and alignment chromatic color, achromatic color and shape similarity (Kurylo, 1997; Schulz, 2001; Schulz, 2002; Schulz et al., 2001; Schulz & Sanocki, 2003). From these results, we can conclude that spatial scenes can be interpreted in more than one way, depending on viewing duration. The following will be a discussion of various lines of empirical evidence that suggest a time course of perceptual organization. Specific emphasis will be placed on the time course of Gestalt grouping as it relates to the proposed line of research.

*Size perception.* In the late 1950s, Gulick and Stake (1957) began by investigating how viewing durations of spatial scenes can affect size perception. Participants observed two triangles that were presented in a 130-foot tunnel. One of the triangles remained in a constant depth location on each trial. The second triangle was adjusted in depth, from 20 to 80 feet, on each trial. Due to the fact that the second triangle was presented at a variety of depth locations, the retinal image of this triangle varied across trials. Although the size of this triangle remained constant, participants were informed that the size of the second triangle would vary. The participants then signaled the experimenter when the two triangles appeared to be the same size. Participants viewed each scene for 0.1 sec, 0.8 sec, or 4.0 sec. An episcotister was the device used to limit exposure to each scene. The experimenters hypothesized that shorter exposure times to the objects in the tunnel may alter the perception of size-constancy when accommodation, convergence, and retinal disparity were the only depth cues available.

Gulick and Stake (1957) found that participants who saw the perceptual scene for
4 seconds were able to make an accurate estimation about the size of the triangles. That is, this estimation was similar to estimations made in unlimited viewing conditions. In this long 4-second condition, size-constancy was therefore perceived. However, the mean estimates of triangle size at 30 to 80 feet in the 0.1 sec exposure condition were significantly different from the 0.8 sec and 4.0 sec condition. It was presumed that in these short durations, size constancy was not maintained. In sum, participants perceived the triangles predominantly by pre-constancy size at 0.1 sec stimulus durations and predominantly by post-constancy size at 4.0 sec durations. Thus, the perception of the size of an object varied with the amount of exposure time that participants were given to view that object.

The findings of Gulick and Stake (1957) suggest a time course of size constancy. More importantly, however, by matching the triangles by size, participants were effectively making grouping assignments at varied stimulus exposure durations. In these experiments, grouping assignments by size changed across time. Without intending to, Gulick and Stake had presented first time course of perceptual grouping evidence using very limited technology, the episcotister.

*Perception of global and local information.* A few years after Gulick and Stake (1957), and with the advent of modern computer technology, the emergence of an organized percept was again analyzed, this time more deliberately and at a finer grain. Navon (1981) studied the initial construction of spatial scenes by global and local processing. To do so, he presented participants with stimuli composed of both global and
local letters, as illustrated by Figure 12. In each stimulus, a cluster of small letters (local letters) was configured to form the shape of a much larger alternative letter (global letter). Participants were asked to identify global letters in some trials and local letters in other trials. To study how global and local letters were initially processed, Navon limited the stimulus duration to 150ms. Participant responses and reaction times were recorded.

Figure 12. Experimental stimuli with global and local letters


Navon (1981) predicted that at stimulus durations as short as 150ms, it would be easier for participants to identify the global letter, as opposed to local letters. Navon found that for correct responses, participant reaction times for naming the global letters in each stimulus was significantly faster than that for the local letters. This suggests an advantage for the processing of global information for object identification early in visual processing.

Other researchers extended Navon’s (1981) research to examine global and local grouping across time. Sanocki (1993) hypothesized that the relative contribution of global and local information to the eventual interpretation of an object could change across time. Specifically, Sanocki predicted that global information would have a greater influence on object identification early in vision and that local information would have a greater
influence on object identification later in vision. To test this hypothesis, Sanocki
presented participants with two stimuli in one of two different orders. Either the global or
local prime preceded a complete target or followed this target. In both sequences, primes
and targets were each shown for a brief 67ms exposure duration. A global or local prime
depicted the exterior or interior features of the complete target, respectively. As shown in
Targets were complete objects that contained the combined global and local properties of
the primes. After the prime target or target prime stimulus sequence was presented,
participants saw four objects and chose one that had been previously displayed as a
target. Response accuracy was recorded.

When global primes were presented before complete targets, the primes were 6%
more accurate than local primes at helping participants choose the correct target.
However, when global primes were presented after complete targets, the primes were 9 %
less accurate than local primes. These results suggest that global information contributed
more to object identification earlier in processing and much less so in later processing.
Conversely, local information contributed more to object identification in later processing
and less so in earlier processing. This evidence suggests a time course of global to local
processing for object identification.

Perception of coarse and fine scene properties. Another way of assessing spatial
scenes is by use of high and low resolution spatial scales, also referred to as spatial
frequency channels or filters. Spatial frequency channels serve to filter visual information
by processing only a restricted range of information (Morrison & Schyns, 2001). Spatial
frequency channels at specific ranges help detect specific types of visual information (e.g., Goffaux, Gautheir, & Rossion, 2003; Nasanen, 1999). Specifically, low spatial frequencies (LSFs) encode coarse scene information (defined by larger less detailed parts) whereas high spatial frequencies (HSFs) encode fine scene information (defined by detailed parts) (Morrison & Schyns, 2001).

Recent research suggests that there is a time course of coarse and fine spatial frequency processing. There are two main views on how this time course occurs. These views are known as the Fixed and Flexible Usage hypotheses. According to the most popular version of the Fixed Usage hypothesis, processing of spatial scales begins with coarse information and then proceeds to fine scene information (Breitmeyer, 1984; Fiorentini, Maffei, & Sandini, 1983; Parker & Costen, 1999; Schyns and Oliva, 1994; Vannucci, Viggiano, & Argenti, 2001). That is, interpretations of spatial scenes are first based on low spatial frequency information and then are later based on high frequency information.

Support for the Fixed Usage hypothesis has been found in several experiments. In one such experiment, Schyns and Oliva (1994) tested human perception of coarse and fine scene details. Participants were presented with ambiguous experimental primes that were hybrids of two perceptual scenes. One perceptual scene was presented as coarse detailed, low frequency information. The second perceptual scene was presented as fine detailed, high frequency information. Primes were presented for 30 and 150ms durations. Target stimuli either matched the coarse blobs or fine edges of the ambiguous hybrid
prime. Participants were asked to determine whether or not each target matched its preceding prime. It was hypothesized that at short prime durations, targets would appear to match the coarse blob interpretation of the ambiguous hybrid prime. Conversely, at long prime durations, it was hypothesized that targets would appear to match the fine edge interpretation of the ambiguous prime.

Schyns and Oliva (1994) found that at short prime durations, participants preferentially matched the course blob interpretation of the primes with the targets. However, in long prime durations, participants preferentially matched the fine edges interpretation of the primes with the targets. These results suggest that in limited exposure durations to spatial scenes, humans tend to organize displays by coarse scene information. However, given longer exposure durations and therefore more processing time to spatial scenes, humans tend to organize displays by fine scene edges. Both the research by Seyns and Oliva (1994) and that of Vannucci et al. (2001) lend support to the Fixed Usage, course to fine hypothesis.

As opposed to the Fixed Usage hypothesis, the Flexible Usage hypothesis suggests that processing can begin with either scale. According to the Flexible Usage hypothesis, visual processing begins with the spatial scale that is most useful (Morrison & Schyns, 2001; Oliva and Schyns, 1997; Schyns and Oliva, 1999). That is, visual processing can begin with high spatial frequency information then followed by low spatial frequency information or begin by low spatial frequency information then followed by high frequency information. The order of processing is believed to rely on task diagnosticity (Oliva & Schyns, 1997).
Recent support for the Flexible Usage hypothesis has been found by Oliva and Schyns (1997). These researchers began by questioning why the human visual system would process coarse scene details before fine details, as indicated by Schyns and Oliva 1994. One explanation could be that coarse scene details are available to the visual system first. To test this hypothesis, Oliva and Schyns (1997) expanded upon their earlier research to determine whether both coarse and fine scene scales are available early in vision or whether one type of scale precedes the other. These researchers presented hybrid primes for 30 ms durations. Targets were either the coarse or fine scale interpretation of those primes. Participants were asked to name the scenes within the targets as quickly and accurately as possible. It was expected that if coarse scene details were available before fine scene details in visual processing, then reaction times to coarse scene targets would be significantly faster than that to fine scene details targets. No significant difference was found between reaction time to coarse and fine scene targets. These results were taken to suggest that both perceptual scales may be available early in vision.

If processing of coarse to fine details cannot be explained by one scale being available before another, what could the explanation be? Oliva and Schyns (1997) hypothesized that perhaps both coarse and fine scales are available in early vision and humans process the scale that works best with their given task. To test this idea, Oliva and Schyns conducted a second experiment in which subjects were presented with practice trials that either required them to categorize spatial scenes by fine or coarse scale
information.

Stimuli presented during the practice trials were unambiguous as to interpretation, either by coarse or fine spatial scale, thus familiarizing participants with one type of scale interpretation task. Immediately after the practice trials, participants were asked to perform the same task with ambiguous coarse/fine stimuli. If scale selection were based on task demands, participants who practiced identifying coarse scene details in the unambiguous practice stimuli would likely perceive the coarse scenes in the ambiguous experimental stimuli. Likewise, participants who practiced identifying fine scene scales in the unambiguous practice stimuli would perceive the fine scenes in the ambiguous experimental stimuli. Indeed this was the case. In short stimulus durations, participants were selecting scales to perceptually organize based on task demands. In effect, perceptual organization was dependent upon what scale of information needed to be processed first, based on the task. The results of both experiments by Oliva and Schyns (1997) support the Flexible Usage hypothesis.

The Fixed and Flexible Usage hypotheses are each interesting interpretations of visual processing by spatial scale. While it is not clear which of these two hypotheses is correct, what is clear is that low and high spatial frequency scales are used in a temporal sequence in visual processing.

Object completion. While some researchers chose to study the time course of perceptual organization in terms of coarse and fine or global and local scene properties, others chose to look at this problem in terms of object completion. For example, Sekuler and Palmer (1992) used matched priming to determine whether participants perceived an
occluded shape as being incomplete or complete at varied prime durations. Participants were briefly presented with ambiguous experimental primes that contained a partially occluded shape. The occluded shape could be perceived as incomplete if amodal completion had not been accomplished or as complete if amodal completion had been accomplished. Prime durations varied from 50 to 400ms. The primes were followed by experimental targets that contained a pair of shapes. The shapes corresponded to either the incomplete interpretation of the prime, the complete interpretation of the prime, or a mix between incomplete and complete prime interpretations. The participants’ task was to determine whether or not the pair of shapes in the target were identical. In this paradigm, participants should have been faster at assessing targets that were similar to the prime. Since the prime was ambiguous as to how it could be interpreted, participant reaction times to the targets shed some light on how they interpreted the prime. If reaction times were faster when assessing incomplete shapes in the target, as compared to complete shapes, then it can be assumed that participants perceived an incomplete shape in the prime. If reaction times are faster when assessing complete shapes in the target, then it can be assumed that participants perceived a complete shape in the prime.

Sekular and Palmer (1992) found that in the shortest prime durations, there was no difference in responses to matching circles or notched circles. However, as prime duration increased, there was a gradual increase in the difference between matching circles or notched targets. Specifically, participants were responding much faster to targets that contained complete circle matches. These results suggest a gradual increase in the interpretation of the occluded prime as a compete shape across time. The perception
of the complete shape in the ambiguous prime emerged across time.

Visual search is another effective paradigm that can be used to test the time course of object completion. Rauschenberger and Yantis (2001) used a visual search task to test the organization of incomplete versus complete shapes across time. These researchers showed that pre-shape completion information can affect performance in short stimulus durations. To do this, these researchers disrupted the perceptual process at different times. Participants were asked to search for a notched disk target in displays that contained distracters that were comprised of paired complete disks and squares. In the control condition, displays contained a notched disk target separated from a nearby square. In the experimental condition, displays contained a notched disk target adjacent to a square occluder. Rauschenberger & Yantis expected that search for the separate notched disk target would be efficient, regardless of the number of distracter square disk elements. If amodal completion was not accomplished in experimental conditions, participants would perceive the notched disk and search for this disk should have been efficient. In the adjacent condition however, after amodal completion, the nonadjacent and notched disk would appear to be a full disk. Therefore, following completion, search for the target in the adjacent condition might be inefficient because of its high degree of similarity to the disk and square distracters. Each display was masked to limit exposure time to 100ms or 250ms. These stimulus durations were selected because amodal completion was thought to occur after 200ms of exposure (Sekuler & Palmer, 1992). When displays were presented for 100ms, search for the notched disk target was efficient in both experimental and control conditions. However, at 250ms stimulus durations,
search for the target in the experimental condition was inefficient. These results show that participants could use pre-amodal completion information when displays were presented for 100ms, but not when displays were presented for 200ms. This research was recently further supported by fMRI evidence showing an incomplete, fragmented interpretation of the prime at 100ms prime durations and an ambiguous interpretation of both the completed shape and initial fragmented interpretation of the prime at 250ms prime durations (Liu, Rauschenberger, Slotnick & Yantis, 2003).

The study of amodal object completion is in effect the study of a type of object constancy, whereby the pre-constancy object image is the incomplete shape and the post-constancy object image is the complete shape. If human perception of constancy can change within time, it is reasonable to assume that human perception of perceptual constancy in other domains might change across time as well. One such type of domain is that of lightness and transparent filters.

*Perception of lightness and transparent filters.* Moore and Brown (2001) showed that visual search processes do not simply involve post-constancy reflectance-based information, but also can reflect pre-constancy retinal-based information. Participants were shown displays that contained gray squares. A filter was placed in front of half of each display so that any square that fell beneath it was reduced in luminance by 50%. Experimental displays contained a light or dark square among medium distracters. The light or dark gray square was the target. The target could appear inside or outside the filtered area in one of 36 randomly selected locations on the stimulus. There were two types of experimental displays: luminance-matched and luminance-unmatched. In
luminance-matched displays, pre-constancy information was relatively ambiguous with respect to whether the target was present or absent. These displays were ambiguous because the target had the same luminance as many of the distracters in the stimulus. In luminance-unmatched displays, however, both the pre-constancy and post-constancy information were unambiguous with respect to target presence. These displays were unambiguous because the targets’ luminance did not match that of any of the distracters. Participants were asked to look at the displays and report whether or not they saw a lighter or darker square, relative to the medium distracters. If pre-constancy information influenced visual search, reaction times should have been longer for the luminance-matched trials. The results showed that reaction times were in fact longer for luminance-matched trials. This suggests that pre-constancy features can affect performance in visual search tasks.

Given that pre-constancy features can affect visual search, it is not unreasonable to assume that these features might affect spatial grouping as well. Recall that Gulick and Stake (1975) unintentionally generated this type of result with their study of grouping by size constancy. Do other types of perceptual grouping operate along a time course? In what follows, several experiments that have tested the time course of perceptual grouping are described in detail.

*Grouping by proximity and alignment.* Do the Gestalt principles of grouping require time to manifest? If so, do some principles manifest more rapidly than others? Kurylo (1997) spoke to these questions by examining the time course of grouping by the Gestalt principle of proximity and by alignment. Note that alignment is an instance of the
Gestalt principle of good continuation whereby elements that are seen as smooth and continuous as perceived as going together (Palmer, 1992). Kurylo asked participants to make decisions about the grouping of elements in masked stimuli with durations of 33 to 150ms exposure times. Experimental stimuli contained black dots that grouped by proximity or alignment in vertical or horizontal orientations. Kurylo found that participants required an 88 ms exposure to group the dots by proximity and a 119 ms exposure to group the dots by alignment.

The results of Kurylo’s research indicate that grouping by the Gestalt principles of proximity and good continuation (alignment) require time. They also demonstrate that grouping by proximity can take less time than grouping by good continuation. These findings are exciting because they suggest that there exists a time course of perceptual grouping. However, Kurylo’s research only tested the time course of grouping relative to the principles of proximity and good continuation. Can the observed time course extend to other types of grouping principles as well?

*Grouping by chromatic color similarity.* Schulz and Sanocki (2003) examined the time course of grouping by chromatic color similarity by presenting ambiguous color stimuli to participants for short and long exposure durations (Figure 13A). Grouping by chromatic color was chosen because of accumulating fMRI evidence suggesting that color is processed along a temporal visual path in which variant retinal wavelength is processed early (in areas V1 and V2) and invariant color constancy is processed later (in area V4) in vision (Zeki, Aglioti, McKeefry, & Berlucchi, 1999; Zeki & Marini, 1998).
Figure 13. Ambiguous chromatic color grouping stimuli and results

![Figure 13](image)

Note. Stimulus and results by Schulz and Sanocki (2003) (a) Ambiguous experimental stimulus. Central column of circles is the pre-constancy, retinal match of the columns of circles on the left and the post-constancy, reflectance match of the columns of circles on the right. (b) Mean percent of grouping by post-constancy reflectance spectrum across time. Standard error indicated by bars.

Experimental stimuli contained five columns of colored circles. Participants were asked to group the central column of circles with the columns on the right or the left hand side by color similarity. The central column of circles in each stimulus was occluded by a tinted transparency. Due to the presence of the transparency, the apparent color of the central circles differed from the actual color. This modified color was referred to as the pre-constancy color. The pre-constancy color of the central column of circles matched the color of the columns of circles on one side of each stimulus. For example, in the stimulus depicted in Figure 13A, the pre-constancy color of the central circles is blue and it matches the color of the circles on the left. However, once the color of the transparency occluding the central column of circles is accounted for, such that the color of the central circles remains, the actual color of the central column of circles can be seen. This is the process of perceptual color constancy. This actual color is referred to as the post-constancy color. The post-constancy color of the central column of circles matched the
columns of circles on the opposing side of each stimulus. For example, in the stimulus depicted in Figure 13A, the post-constancy color of the central circles is red and matches that of the columns of circles on the right. Due to the pre- and post-constancy color interpretations of the central circles, the experimental stimuli presented by Schulz and Sanocki were ambiguous as to how the central column of circles could be grouped by color similarity.

Masks followed each stimulus to limit exposure time to 200, 500, 1100, and 2000ms. Participants directly reported the grouping of the central column of circles by key press. As shown in Figure 13B, Schulz and Sanocki (2003) found that grouping was primarily based on pre-constancy color when the stimuli were presented for brief exposure durations (pre-constancy 88%, post-constancy 12% at 200ms). Alternatively, grouping was primarily based on post-constancy color in the longest stimulus duration (pre-constancy 18%, post-constancy 82% at 2000ms). In sum, grouping began by one color and then gradually changed to occur by an alternative color over time.

The finding that pre-constancy color information is more salient in short stimulus durations has raised some interesting questions. One such question is whether pre-constancy information can be used to speed grouping responses in unlimited exposure conditions. Schulz and Sanocki (2003) conducted a second experiment to speak to this question. Participants were presented with experimental stimuli that were identical to those presented in Experiment 1 (see Figure 13A). However, stimuli in the second experiment were presented for unlimited exposure durations and reaction times were recorded. The difference between pre- and post-constancy color was explained to
participants. Participants were then asked to group by the pre-constancy color of the central column of circles in some trial blocks and by the post-constancy color in other trial blocks. Since pre-constancy color grouping was observed to be more salient in short stimulus durations in the Experiment 1, it was expected that participants would be faster when grouping by pre-constancy color, relative to grouping by post-constancy color.

The mean reaction time for correct pre-constancy grouping responses was 605 ms and that for post-constancy grouping responses was 175 ms greater. Thus, pre-constancy color grouping was shown to be considerably faster than post-constancy grouping. In a third experiment to further test this idea, Schulz and Sanocki (2003) presented the stimuli for an unlimited amount of time. Participants were asked to group the central column of circles by similarity and responses and reaction times were recorded. Based upon their responses, the participants were divided into three groups, pre-constancy groupers, post-constancy groupers, and mixed groupers. The mean reaction time for each of these types was then computed. Schulz and Sanocki found that the participants who grouped predominantly by pre-constancy color had a mean reaction time of 708ms, whereas those who grouped predominantly by post-constancy color had a mean reaction time of 1600ms. Thus, faster responding was associated with pre-constancy color grouping and slower responding with post-constancy grouping.

The results of Schulz and Sanocki (2003) show that color similarity grouping can occur in more than one way, depending on exposure time. Exposure time limits may be imposed by the task, as in Experiment 1, or imposed by the user, as in Experiment 2. These findings also show that knowledge about the type of grouping that is available in
short stimulus durations can be used to predict the relative speed of response times based on that type of grouping.

It is important to note that Schulz and Sanocki (2003) asked participants to explicitly report the grouping of color stimuli. In these conditions, it is hard to evaluate the extent to which direct reports assess what participants are really seeing, as opposed to what they think they should be seeing. In addition, participants in these experiments were required to choose either the pre- or post-constancy match of the central column of circles, with no way to indicate if they wished to choose both.

To speak to this issue, priming was used to determine whether color similarity grouping operated differentially on pre- and post-constancy information as a function of prime duration (Schulz, 2002; Schulz et al., 2001). Priming was chosen as a more indirect measure. Participants were presented with ambiguous prime displays, which were identical to the experimental stimuli used by Schulz and Sanocki (2003) (Figure 13A). Participants did not respond to the primes. A mask was used to limit prime duration to 25 to 1750 ms. The masks were followed by unambiguous target displays that matched either the pre- or post-constancy grouping solution of the prime (see Figure 14A and 26B). For example, with respect to the stimulus presented in Figure 13A, Figure 14A is the pre-constancy grouping solution. Alternatively, Figure 14B is the post-constancy grouping solution. Participants reported whether the central circles of the unambiguous target stimuli grouped with the circles on the right or left. Responses and reaction times were measured. Reaction times to pre- and post-constancy match targets were compared to target controls. Based on Schulz and Sanocki's (2003) results, it was expected that
reaction times would be faster that controls when targets matched the pre-constancy grouping solution of the prime at short prime durations. Similarly, it was expected that reaction times would be faster that controls when targets matched the post-constancy grouping solution of the prime at long prime durations.

Figure 14. Unambiguous color grouping targets

Note. (a) Pre-constancy match experimental target. (b) Example post-constancy experimental match target.

The results revealed that reaction times to pre-constancy match targets were significantly faster than those to post-constancy match targets at prime durations shorter than 450ms (see Figure 15A). This finding was consistent with Schulz and Sanocki's (2003) direct report research. However, reaction times to post-constancy match targets were not significantly different from zero at prime durations beyond 450ms (see Figure 15B). This finding was inconsistent with previous direct report research and has been taken to suggest the possibility that there exists an ambiguity between pre- and post-constancy color representations beyond 450ms.
**Figure 15. Results for ambiguous priming study**

![Diagram A](image1.png) ![Diagram B](image2.png)

*Note.* Results presented by Schulz et al., 2001. Standard error indicated by bars. (a) Mean pre- and post-constancy match advantage by prime durations in between subjects condition. (b) Mean pre- and post-constancy match advantage in within-subjects condition.

In sum, the previously detailed studies suggest that there exists a time course of color similarity grouping in that observers can switch from pre- to post-constancy color information with the increase in stimulus duration. But what about other types of similarity grouping like lightness, shape, texture, and pattern? Does the observed time course of grouping extend to these attributes as well?

*Grouping by achromatic lightness.* To test the generality of the time course of similarity grouping, Schulz (2001) tested grouping by lightness and shape similarity at short and long stimulus durations. In a first experiment, ambiguous lightness stimuli were presented for 200 or 2000ms. Experimental stimuli (see Figure 16A) followed luminescence constancy stimuli used in prior research (Rock, Nijhawan, Palmer & Tudor, 1992). Experimental stimuli contained 5 columns of circles. Participants were asked to group the central column of circles with the columns on the right or the left hand side by
lightness similarity and responses were recorded. The critical central column of circles was manipulated by depicting a tinted transparency in front of it. As a result, the central circles matched the pre-constancy lightness of the two columns on one side of the stimulus. If participants grouped the central circles by similarity of pre-constancy lightness, then they would group it with this side of columns. In Figure 16A for example, the central column of circles matches the pre-constancy lightness of the circles on the left hand side. However, after the transparency has been accounted for, the actual post-constancy lightness of the central circles can be seen. The post-constancy lightness of the central circles matched the opposing two columns of circles in each stimulus. If participants grouped the central circles by similarity of post-constancy lightness, they would group it with these opposing columns. In Figure 16A for example, the post-constancy lightness of the central circles matches the circles in the columns on the right hand side.

Figure 16. Ambiguous achromatic color grouping stimuli and results

Note. (a) Experimental stimulus, Schulz (2001). Central circles match pre-constancy lightness of columns on left and post-constancy lightness of columns on right. (b) Mean percent of grouping by post-constancy lightness across time.
As shown in Figure 16B, for the lightness similarity experiment, grouping was based primarily on pre-constancy lightness at the shorter exposure time (80.5% pre-constancy, 19.5 post-constancy at 200ms). At the longest exposure time, however, pre-constancy grouping decreased and there was an increase in post-constancy grouping (38% pre-constancy, 62% post-constancy at 2000ms). These results show that grouping can begin by pre-constancy lightness and then switch to occur by post-constancy lightness when exposure time is increased.

*Grouping by shape similarity.* In a second experiment by Schulz and Sanocki (2001), ambiguous shape stimuli were presented for 200 or 2000ms. Experimental stimuli (Figure 17A) followed shape completion stimuli used in prior research (Palmer, Neff & Beck, 1996). Stimuli contained 5 columns of shapes. Participants were asked to group the central column of shapes with the columns on the right or left hand side by similarity and responses were recorded. Two of the outer columns on each experimental stimulus were composed of half circles. The opposing two outer columns were composed of full circles. The critical central column was composed of circles. However, the central column was manipulated by an opaque strip that was depicted to be in front of it. As a result, the incomplete shape of the central circles appeared to match the shape of the two columns of incomplete circles on one side of the stimulus. If participants grouped the central circles by similarity of incomplete shape, then they would group the central circles with this side of columns. In Figure 17A, for example, the central circles match the incomplete shape of the circles on the left. However, after the opaque strip is accounted for, the complete shape of the central circles should be perceived. This
interpretation should match the shape of the opposing columns of circles. If participants
grouped the central circles by similarity of complete shape, they would group the central
circles with these opposing columns. In Figure 17A, for example, the central circles
group by complete shape with the columns of shapes on the right.

Figure 17. Ambiguous shape grouping stimuli and results

Note. (a) Experimental stimulus, Schulz (2001). Central column of shapes matches incomplete shape of columns of left
and complete shape of columns on right. (b) Mean percent of grouping by complete shape across time.

The mean percentages of grouping experimental stimuli by complete shape, for
each stimulus duration, are plotted in Figure 17B. Grouping by incomplete shape was
predominant at the shortest exposure time (58% incomplete shape, 42% complete shape
at 200ms). At the longest exposure time, however, grouping switched to occur
predominantly by complete shape (25% incomplete, 75% complete). These results
suggest that grouping can begin by incomplete shape and then switch to occur by
complete shape as exposure time increases.
In review these findings, it becomes clear that exposure time is a critical factor to consider when determining the perceived organization of spatial displays. It is important to note that with the exception of the stimuli of Gulick and Stake (1957), all of the stimuli detailed in the previously mentioned time course experiments were presented on computer monitors. The results of these experiments should therefore bear some information on how humans perceive and group elements presented on computerized interface displays. Further, research by Bruno, Domini and Bertamini (1997) suggests that varying exposure durations to pictorially presented displays (such as computerized displays) can have a much different effect on organization than varying exposure duration to displays that include binocular parallax (such as real world scenes). Bruno et al. compared the time course of amodal completion in pictorial and binocular parallax displays. This research followed research conducted by Sekuler and Palmer (1992) who found evidence for a pre-completion interpretation (i.e., incomplete shape) of pictorially presented amodal completion displays when stimuli were shown for less than 200ms. When presenting displays pictorially, Bruno et al. found evidence for the perception of a pre-completion interpretation of the stimuli at 100ms stimulus durations. This result was consistent with the results obtained by Sekuler and Palmer. However, inconsistent with previous pictorial research, the perception of a pre-completion interpretation of the stimuli was not found when similar displays were presented under the condition of binocular parallax. Bruno et al. concluded that Sekuler and Palmer’s early pre-completion result might have been obtained because pictorial displays were used. While this finding may have raised questions about the generality of the Sekular and Palmers’ result in real
world spatial displays, it also heightens curiosity about the nature of time limited perceptual organization in computerized displays. Given the findings about exposure time limits to user interface and the recent findings in studies of the time course reviewed above, it seems logical to think that the time course of perceptual grouping should be considered when designing user interface displays.

**Current line of inquiry**

What users perceive in time limited conditions may be more important, more salient, and more usable than what designers perceive in unlimited exposure conditions. Time limited task demands may call for a modification in user interface design, one that focuses on recent discoveries in perceptual organization and, specifically, perceptual grouping. Researchers in user interface design have found effects of grouping in unlimited time exposure conditions that have guided the organization of buttons, pictures, icons and text in user interface design (Card, 1982; Niemela & Saarien, 2000; Tullus, 1981; Tullis, 1984; Tullus, 1986). These findings have paved the way for the development of grouping and interface design principles (Bellcore, 1995; Bailey, 1982; Card, 1982; Danchak, 1976; Galitz, 1985; Holden, Adolf & Williges & Williges, 1981; Jones & Okey, 1997; Moore & Fitz, 1993; Stewart, 1976; Streveler & Wasserman, 1984; Tullis, 1983; Tullis, 1988; Woolford, 1997). However, researchers in vision science have found that grouping can occur in more than one way across time (Kurylo, 1997; Schulz, 2001; Schulz, 2002; Schulz et al., 2001; Schulz & Sanocki, 2003). These findings lead us to question whether grouping can be perceived differentially at short and long durations.
in user interface display layouts.

It would be interesting to determine whether grouping operates differentially across time in user interface displays because display layouts are often ambiguous with respect to organization. For example, in the website shown in Figure 18, there are a variety of grouping principles that can affect the users’ perception of the grouping of the navigational pushbuttons. If organizing by similarity and common region, the user may perceive the pushbuttons to form one large group, as the pushbuttons are all depicted in a similar pattern and are all located on one common dark region. If organizing by alignment, the user may perceive the pushbuttons to form two groups, one group in a top row and a second group in a bottom row, as the pushbuttons appear to be aligned in two rows. If organizing by proximity, the user may group the bottom central pushbutton and the ‘Email Webmaster’ link. Hence, the pushbuttons in Figure 18 are ambiguous with respect to grouping. While this website serves as one example, there are countless other examples of ambiguous display groupings in both software applications and websites.

Figure 18. Ambiguous website

Note. Pushbuttons can be organized by several types of Gestalt principles of grouping.
Given ambiguous grouping displays such as the one in Figure 18 and given what is known about the time course of grouping in visual perception, it is logical to question which grouping principles are more salient in interface displays at various points in time. Addy (2000) began to address this question with his research, but his results were limited to unlimited exposure time conditions. Perhaps it is too simplistic to think that some grouping principles are always necessarily dominant over others in interface displays. Rather, perhaps grouping principles can be manipulated by other factors to become dominant over others in certain conditions. Further, perhaps these factors influence grouping differentially at particular stimulus exposure durations.

There are likely to be a variety of factors that influence which grouping principles are perceived as dominant at various points in time. One such factor may be the extent to which each grouping principle is presented as being global or local, relative to other grouping principles in an interface display. Recall that according to research on global and local properties of spatial displays, global properties have been shown to be perceived before local properties across exposure time (e.g., Navon, 1981; Sanocki, 1993). It would be interesting to determine if this discovery in vision extends to the Gestalt principles of grouping and specifically whether it applies to these principles when presented in user interface displays.

To test the time course of global and local Gestalt grouping in interface displays, Schulz and Sanocki (2002) initially conducted three pilot experiments. Each of the three pilot experiments utilized the same design, with stimuli varying slightly. Participants
were asked to make decisions about experimental stimuli, displayed on a computer monitor, in which a central pushbutton was ambiguous as to how it could be grouped (see Figure 19A). As shown in Figure 19A, participants could group the central pushbutton with one side of pushbuttons by a global grouping principle (in this experiment, color similarity), or with the opposing side of pushbuttons by a local grouping principle (in this experiment, connectedness). Color similarity was considered to be more global than connectedness because this grouping principle was the largest scale feature within these stimuli. Unambiguous control stimuli, which had only one correct grouping solution, were also presented to make certain that participants were actually grouping the central pushbutton. Experimental stimuli were shown for 200 and 2000ms. Responses and reaction times were recorded.

Figure 19. Pilot stimulus 1 and results

Note. Stimulus and results by Schulz and Sanocki (2002) (a) Ambiguous stimulus in first pilot experiment. Central pushbutton groups with pushbuttons on left by global color similarity and with pushbuttons on right by local connectedness. (b) Results from first pilot experiment. Graph shows number of participants who grouped predominantly by color similarity, connectedness, or mixed properties by exposure duration.

It was hypothesized that at short stimulus durations, participants would group stimuli predominantly by color similarity, the more global scene detail. At long stimulus
durations, however, it was predicted that grouping could occur by both color similarity and connectedness, the local scene detail. This hypothesis was based on recent research in vision science that has shown a time course of global to local perceptual processing (Kimchi, 1998; Kimchi & Hadad, 2002; Sanocki, 1993; Sanocki, 2001).

Once the data were collected, participants were classified as either color similarity groupers, mixed groupers, or connectedness groupers. Figure 19B depicts these findings. As shown, most participants in the short stimulus duration grouped by the global principle of color similarity (13 by color similarity, 2 by connectedness in the 200ms condition). At the long stimulus duration, however, some participants grouped by local principle of connectedness and some participants grouped by a mixture of both principles (8 by color similarity, 4 by mixed principles, and 3 by connectedness in the 2000ms condition). In summary, grouping began predominantly by a global Gestalt principle of color similarity but then switched to occur by both global and local principles. The results of the first pilot study suggest the potential for a time course of global to local Gestalt grouping in interface displays.

One of the shortcomings of the stimuli presented in the first pilot experiment was the use of rectangular buttons. Schulz and Sanocki (2002) questioned whether the rectangular buttons in the stimuli grouped by connectedness in the horizontal plane in the same way that they grouped by connectedness in the vertical plane. If not, this may have affected horizontal grouping by connectedness of the central pushbutton with the vertical columns of pushbuttons on either side. To address this issue, pushbuttons presented in the second pilot experiment were depicted as squares instead of rectangles. In addition, the
columns of pushbuttons on either side of the central pushbutton were reduced to a single button, rather than columns, so that participants would only group by horizontal connectedness. Figure 20A is an example of the modified stimuli presented in the second pilot experiment.

Figure 20. Pilot stimulus 2 and results

Note. Stimulus and results by Schulz and Sanocki (2002) (a) Ambiguous stimulus used in second pilot experiment. A central pushbutton groups with pushbutton on left by global color similarity and with pushbutton on right by local connectedness. (b) Results for second pilot experiment. Graph shows number of participants who grouped predominantly by color similarity, connectedness, or mixed properties by exposure duration.

The second pilot experiment was conducted with the modified stimuli. This second pilot experiment revealed a trend similar to the first pilot experiment, as seen in Figure 20B. Participants began by grouping predominantly by global Gestalt properties (color) in short stimulus durations. However, at longer durations, participants grouped by a mixture of global (color) and local (connectedness) properties.

The first and second pilot experiments each showed a moderate global to local grouping effect as exposure time increased. Upon review of these results, Schulz and Sanocki (2002) questioned whether a global to local grouping result could be found if the
Gestalt principles representing the global and local cases were switched. That is, these researchers questioned whether their results could be replicated if color were depicted as being less global and connectedness were depicted as being less local. To make color less global, Schulz and Sanocki (2002) reduced the color of the pushbuttons from bold to pastel. To make connectedness less local, these researchers thickened the lines depicting connectedness and tripled these lines in each stimulus. Figure 21A is an example of the modified stimuli presented in the third pilot experiment.

Figure 21. Pilot stimulus 3 and results

Note. Stimulus and results by Schulz and Sanocki (2002) (a) Ambiguous stimulus in third pilot experiment. A central pushbutton groups with pushbuttons on left by color similarity and with pushbuttons on right by connectedness. (b) Results for third pilot experiment. Graph shows number of participants who grouped predominantly by color similarity, connectedness, or mixed properties by exposure duration.

Figure 21B summarizes the results of the third pilot study. As one can see, the trend found in the first and second pilot experiments has reversed in the third pilot experiment. Grouping was more reliant on the less local Gestalt principle of connectedness in the short stimulus duration. Schulz and Sanocki (2002) reviewed the results of the three pilot experiments and proposed that (1) Gestalt principles of grouping can be depicted as being global or local in user interface displays and (2) that Gestalt
principles that are depicted more globally in interface displays tend to be utilized in shorter stimulus durations.

At the conclusion of the three pilot experiments, Schulz and Sanocki (2002) continued to improve the experimental stimuli. One area of concern was the use of the Gestalt principle of color similarity as the global grouping principle in the stimuli. While color similarity grouping in the displays seemed to be naturally more global than connectedness grouping, it was difficult to operationally define this. How much more global was color similarity when compared to connectedness? Furthermore, color was depicted as being part of the pushbuttons whereas connectedness was depicted as independent of these buttons.

The preceding issues motivated several changes to the design of the experimental stimuli in the present study. First, Schulz and Sanocki (2002) decided that color similarity was not the best grouping principle to use in the experimental stimuli, because it was difficult to define and manipulate the global magnitude of the color. Rather, a new grouping principle of common region was used for the present experimental displays. Common region was chosen because it is a global principle that could be depicted as being independent of the pushbuttons, much like the Gestalt principle of connectedness.

In addition to making some changes to the types of Gestalt principles used in the displays and the global and local depictions of these principles, rules were created to define how Gestalt principles could be classified as global or local scene properties. These rules were used as a guide in stimulus creation. Prior researchers have defined
global scene properties as the largest size shapes of objects in perceptual scenes (Sanocki, 2001; Navon, 1977; Navon, 1981). Global grouping in the current displays is therefore defined as an organizational property that is relatively large in size. Conversely, prior researchers have defined local properties as being more internal, interior and generally smaller shapes in perceptual scenes (Sanocki, 1993). Local grouping in the current displays is therefore defined as an organizational property that is relatively small in size. For example, with respect to Figure 9, the large ‘Altitude’ and ‘Speed’ boxes represent global grouping whereas the small “Status” and “Measure” boxes represent local grouping.

Perhaps the best way to understand the difference between global and local grouping in the present studies is to quantify them. Global grouping defines a quantitative difference in size, relative to a local comparison. In the present experiments, global grouping cannot be defined as global unless it is contrasted with a local grouping equivalent. Due to its quantitative nature, global grouping can vary in degree. Thus, grouping in a particular stimulus can be slightly more global than a local comparison or much more global, depending on the difference in global and local size. Greater differences in size define a more dramatic difference between global and local grouping. It should be noted that there must be some upper limit to global grouping. For example, there must be some instances when global grouping is so large in size that grouping is not perceived at all. While of interest, these cases will not define global grouping in the present study and will not be represented in the present stimulus displays. Likewise, it is possible that there must be some lower limit to local grouping, one in which grouping is
no longer perceived. These cases will not define local grouping in the present studies, nor will they be represented in the present displays.

In the present study, participants were presented with stimuli, displayed on a computer monitor, in which a central pushbutton was ambiguous as to how it could be grouped with the pushbutton to its immediate right and left side. The central pushbutton could group with the pushbutton on one side of the display by a global Gestalt principle and with the pushbutton on the opposing side of the display by a local Gestalt principle. In Figure 22A, for example, the central pushbutton groups by common region with the pushbutton on the left side and groups by connectedness with the pushbutton on the right side. Common region is depicted as being global in size scale when compared to connectedness.

Figure 22. Primes used in reported experiments

Note. (a) Example experimental prime. Central pushbutton groups by common region with pushbuttons on the left and by connectedness with pushbuttons on the right. In this display, common region is global relative to connectedness. (b) Example experimental prime. Central pushbutton groups by common region with the pushbuttons on the left and by connectedness with pushbuttons on the right. (c) Example neutral prime. Central pushbutton does not group with the pushbuttons on the left or right hand side of the display.
Using the previously described type of stimuli, three experiments tested the time course of global and local grouping in interface displays. Experiment 1 was designed to test whether global depictions of Gestalt grouping principles are more salient than local depictions of these principles at short stimulus durations. To address this question, Hypothesis 1 was proposed.

Hypothesis 1: At short exposure durations, users group elements in ambiguous interface displays by global Gestalt grouping principles. Conversely, at longer exposure durations, users can group elements in ambiguous interface displays by both global and local grouping principles.

Another way to determine which grouping principles are first perceived is to determine which principles quicken RT. Experiment 2 was designed to test whether global depictions of Gestalt grouping principles can be grouped significantly faster than local depictions. To address this question, Hypothesis 2 was proposed.

Hypothesis 2: It is hypothesized that users group interface displays by global grouping principles at a faster rate than they group by local grouping principles, when interface displays are presented for unlimited durations.

If the results of Experiments 1 and 2 suggest that particular type of pushbutton grouping is more salient at short stimulus durations than another, this information might be used to redesign elemental organizations in user interface displays. How will these changes be received by users who are familiar with seeing displays with other types of configurations? Specifically, how does knowledge of what has ‘gone together’ in the past
influence what ‘goes together’ in the future in interface displays? As seen in research by Oliva and Schyns (1997), task diagnosticity has been shown to have an effect on perceptual organization by spatial scales. Furthermore, prior knowledge has been shown to have an effect in vision (e.g., Biederman, 1987; Peterson, 1994; Shepard, 1983). Experiment 3 was designed to test whether having prior knowledge of display groupings could affect user efficiency when users must rely on alternative groupings. To address this question, Hypothesis 3 was proposed.

**Hypothesis 3:** It is hypothesized that when users have experience with grouping by particular size scale (global or local) in interface displays, it will be easiest for participants to group by this scale during future exposures to the display.
General Method

Participants

Participants volunteered in exchange for extra credit in undergraduate psychology courses at the University of South Florida. According to self report, all participants had normal or corrected-to-normal vision, and all were native English or bilingual speakers with English as one of their spoken languages. Each volunteer participated in only one of the experiments. Data from participants who did not meet the above criteria were omitted from the analysis.

Design

Across experiments, the following remained the same. The size scale of Gestalt grouping was manipulated as the independent variable. There were two levels for size scale: global and local. Responses and reaction times were recorded as the dependent variables. Experimental sessions lasted no longer than 30 minutes.

In Experiment 1, I compared global and local grouping in a 50ms and 4000ms prime duration. Size scale (global or local) was presented within participants, randomly with replace. The 50ms and 4000ms prime durations were manipulated between participants. Participants were randomly assigned to a prime duration (50ms or 4000).

In Experiments 2A and 2B, I compared reaction time for global and local grouping in unlimited exposure durations. Size scale (global or local) was presented within
participants, randomly with replace. For both Experiments 2A and 2B separately, a repeated measures ANOVA was used to compare reaction times for the global and local conditions.

In Experiment 3, I compared size scale (global and local) of grouping in a test period that immediately followed a global or local training session. Participants were randomly assigned to one condition.

**Stimuli**

Each experiment used the same types of stimuli. Primes and target stimuli were created with Power Point, a presentation program. Ready signals and masks were created with Adobe Photoshop, a digital imaging program. Stimuli were saved with an 8-bit pixel depth. Stimuli were 640 pixel width by 480 pixel height and displayed in the center of the screen. Stimuli were displayed on a Macintosh G3 PowerBook. A ready signal preceded the control, experimental and practice stimuli. A mask immediately then followed each stimulus after it had been presented. In Experiment 1, a target followed the mask.

Ready signal. The ready signal consisted of a white screen with a black plus sign in the center. The ready signal prompted the participants to focus their attention on the fixation cross where stimuli subsequently appeared.

Primes. In the experimental primes, a central pushbutton was depicted with one pushbutton to its right and one pushbutton to its left (e.g., Figure 22A). The central pushbutton grouped with the pushbutton on one side by a global Gestalt grouping property and with the pushbutton on the opposing side by a local Gestalt grouping
property. Primes were therefore ambiguous with respect to how the central pushbutton should be grouped. Common region and connectedness were the Gestalt grouping properties depicted in the primes.

The color of common region and connectedness was counterbalanced. In one stimulus type, common region was depicted in light gray and connectedness was depicted in dark gray, as seen in Figure 22A. In a second stimulus type, common region was depicted in dark gray and connectedness was depicted in light gray, as seen in Figure 22B. Neutral primes were also created in which a central pushbutton did not group with the pushbuttons on either the right or the left hand side (see Figure 22C).

A mirror reflection of each of the described primes was created so as not to bias right or left side grouping responses. In summary, a total of 4 primes were created (2 achromatic color schemes X 2 mirror reflections).

Targets. Unambiguous target stimuli contained a central pushbutton that could group with a pushbutton on either side by either global or local interpretation of the primes, but not by both. Two target stimuli were created for each prime. One depicted the unambiguous global grouping interpretation of the prime and the other depicted the unambiguous local grouping interpretation. For example, with respect to Figure 22A, Figure 23A is the unambiguous global grouping interpretation and Figure 23B is the unambiguous local grouping interpretation. Because there were four primes and two targets for each prime, there were a total of eight targets.
Figure 23. Targets used in reported experiments

Note. (a) Example unambiguous target. The central pushbutton groups with the pushbutton to the left. Relative to the experimental stimulus that it was modeled after, the central pushbutton groups by global properties. (b) Example unambiguous target. The central pushbutton groups with the pushbutton to the right. Relative to the experimental stimulus that it was modeled after, the central pushbutton groups by local properties.

Practice stimuli. Practice stimuli contained a central pushbutton that could group with either a pushbutton to its right or left by the Gestalt principle of proximity (see Figure 24). Practice stimuli were therefore unambiguous with respect to how the central pushbutton should be grouped. The grouping principle of proximity was chosen, as opposed to other principles, because it is not presented in the prime or target stimuli. Participants therefore learned to group the pushbuttons in these practice displays without being biased toward grouping by a particular grouping principle depicted in the experimental stimuli.

Figure 24. Practice stimuli used in reported experiments

Note. (a) Unambiguous practice stimulus. Central pushbutton groups with pushbutton on the left hand side by proximity. (b) Unambiguous practice stimulus. Central pushbutton groups with pushbutton on the right hand side by proximity.
Mask. A mask was used to disrupt processing after each type of stimulus was shown. The mask consisted of a white background covered with scribbled achromatic lines. The achromatic colors were cloned from the colors in the prime, target, and practice stimuli. The mask was the size of the area of the prime, target and practice stimuli.

Procedure

In each of the experiments, participants were asked to read and sign an informed consent form in order to participate. Participants sat approximately 24 inches from the 13.1 inch Apple monitor. The visual angle, a calculation of the size of the stimulus and its distance from the participant, was held approximately at 24 X 18 degrees. Stimuli were displayed using PsyScope, an interactive presentation program (Cohen, MacWhinney, Flatt & Provost, 1993).

At the start of each trial, PsyScope presented the ready signal. The ready signal was followed by a practice, prime or target stimulus. The order of the stimuli was chosen randomly with replacement. The onset of each stimulus was cued by a brief tone. The tone served as an auditory cue to alert the participants to focus on the stimuli. Following each stimulus, a mask was displayed.

Participants were asked to respond as quickly and accurately as possible. Participants indicated their responses by pressing either one of two keys on a standard computer keyboard. Responses and response times were recorded by PsyScope. Response times that were more than three SD from each participant mean were omitted. Data from
participants that were more than three SDs above or below the group mean were also omitted.
Experiment 1

What type of grouping information do users best perceive in time limited user interface conditions? Is this information the same as that perceived in longer exposures by interface designers? Experiment 1 was designed to address these questions.

Method

In Experiment 1, common region served as the global grouping principle and connectedness served as the local grouping principle. For example, in Figure 22A, the central pushbutton groups by common region with the pushbuttons on the left side and grouped by connectedness with the pushbuttons on the right side. In this stimulus, common region is depicted as being global when compared to connectedness, because it is depicted as being the larger scale shape in area and more of an exterior contour. More specifically, in this stimulus, common region is depicted as being 2.5 times as high and 7 times as wide as connectedness.

Participants were familiarized with the grouping task by grouping unambiguous practice displays for ten trials. After the practice period, ambiguous experimental and neutral primes were randomly presented to participants via PsyScope. There were 12 blocks of trials, each containing 24 trials each. Trial blocks were separated by a rest period that continued until the participant decided to begin the next block. Primes were masked to limit exposure duration to 50 and 4000ms between participants. Note that the short stimulus duration utilized a 50ms exposure duration, instead of the 200ms duration
that was presented in the pilot experiments, because user eye saccades can be as short as 50ms. The 50ms exposure time was therefore more likely to provide participants with visual information that is representative of what they would see in interface displays during time limited tasks. The long exposure duration utilized a 4000ms exposure time, as opposed to the 2000ms exposure time used in the pilot experiments, because it was expected that information perceived in this amount of time more closely approximates what users perceive in unlimited exposure times to interface displays. Targets followed the mask screens and participants responded when they saw the target.

For each trial, participants were asked to observe the primes and unambiguous targets before deciding how to group the central column of pushbuttons in the targets. Figure 25 depicts all stimulus sequences. In Figure 25A, an ambiguous experimental prime is followed by a global (common region) match target. In Figure 25B, an ambiguous experimental prime is followed by a local (connectedness) match target. In Figure 25C, a neutral prime is followed by a global (common region) match target. In Figure 25D, a neutral prime is followed by a local (connectedness) match target. Right and left side grouping responses were recorded for the unambiguous targets, along with reaction times.
Results and discussion

Data from 21 participants were collected for the 50ms condition and from 19 participants for the 4000ms condition. In the 50ms prime duration, I omitted the data
from three participants because one participant did not accurately group over 10% of the unambiguous targets, one participant needed glasses in order to see the computer screen but did not bring glasses to the experiment and one participant reported having vision problem that had not been corrected by glasses or contacts. In the 4000ms prime duration, I omitted the data from one participant because she did not accurately group over 10% of the unambiguous targets. In the 50ms condition, participants ranged in age from 18 to 42 with a mean age of 22.4. In this condition, all of the participants were female. In the 4000ms condition, the participants ranged in age from 18 to 28 with a mean age of 21.6. In this condition, there were 15 female participants and 3 male participants.

For both the 50ms and 4000ms prime durations, the mean reaction times to common region targets that followed experimental primes were subtracted from mean reaction times to common region targets that followed neutral primes. This score was considered to be the global (common region) facilitation score. Likewise, the mean reaction times to connectedness targets that followed experimental primes were subtracted from mean reaction times to connectedness targets that followed neutral primes. This score was considered to be the local (connectedness) facilitation score. Positive scores signified an advantage whereas negative scores signified a disadvantage. The global (common region) and local (connectedness) facilitation scores were compared. A mixed ANOVA (Size scale X Prime duration) was used to compare facilitation reaction times. Figure 26 depicts the results for Experiment 1. There was a main effect of size scale [global facilitation (M = 6.17 ms, SD = 16.56 ms), local facilitation (M = -4.33 ms, SD = 17.62 ms), F (1, 34) = 7.06, p = .012]. This result
suggests that participants perceived global groupings to be more salient than local groupings in both prime durations. There was no main effect of prime duration, $F(1, 34) < 1$. This result suggests that there was no difference in reaction times across the 50 ms and 4000 ms prime durations. There was no interaction of scale and prime duration, $F(1, 34) < 1$. This result suggests that there is no significant difference between global facilitation for the 50ms and 4000ms prime durations. Likewise, this result suggests that there is no significant difference in local facilitation between the 50ms and 4000ms prime durations.

Figure 26. Results for Experiment 1

---

**Note.** Global and local facilitation by prime duration. Standard error indicated by bars.

Figure 27A, depicts a box plot of global and local facilitation means for all conditions in Experiment 1. The box plot provides additional information about the skew, quartiles and median of the distribution of means for each condition. These specific values can be found in Table 1. Additionally, the pictures presented in Figure 27B, C, D, and E depict the distribution of raw global and local grouping reaction times for each prime duration. These figures provide additional information about the distributions. As one can
see, each of the four distributions is very similar because each has a positive skew. However, low reaction time scores are more frequent for global grouping when compared to local grouping across prime durations. This explains why global facilitation was significantly greater than local facilitation in Experiment 1.

Figure 27. Additional analysis for Experiment 1 data
Table 1. Characteristics of RT distribution for Experiment 1

<table>
<thead>
<tr>
<th>Experiment 1</th>
<th>Mean</th>
<th>SD</th>
<th>Skew</th>
<th>5%</th>
<th>25%</th>
<th>50%/Med</th>
<th>75%</th>
<th>95%</th>
<th>S-W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glo Main</td>
<td>6.17</td>
<td>16.57</td>
<td>0.03</td>
<td>-21.70</td>
<td>-5.25</td>
<td>5.50</td>
<td>19.25</td>
<td>33.25</td>
<td>0.853</td>
</tr>
<tr>
<td>Loc Main</td>
<td>-4.33</td>
<td>17.62</td>
<td>-0.22</td>
<td>-43.75</td>
<td>-14.00</td>
<td>-5.50</td>
<td>4.75</td>
<td>28.15</td>
<td>0.494</td>
</tr>
<tr>
<td>50 ms x Glo Int</td>
<td>6.50</td>
<td>11.38</td>
<td>0.62</td>
<td>-12.00</td>
<td>0.50</td>
<td>5.50</td>
<td>10.25</td>
<td></td>
<td>0.306</td>
</tr>
<tr>
<td>50 ms x Loc Int</td>
<td>-1.67</td>
<td>13.72</td>
<td>0.45</td>
<td>-20.00</td>
<td>-14.00</td>
<td>-4.50</td>
<td>7.75</td>
<td></td>
<td>0.295</td>
</tr>
<tr>
<td>4000 ms x Glo Int</td>
<td>5.83</td>
<td>20.86</td>
<td>-0.03</td>
<td>-37.00</td>
<td>-9.00</td>
<td>4.50</td>
<td>23.25</td>
<td></td>
<td>0.985</td>
</tr>
<tr>
<td>4000 ms x Loc Int</td>
<td>-7.00</td>
<td>20.88</td>
<td>-0.17</td>
<td>-48.00</td>
<td>-16.75</td>
<td>-7.00</td>
<td>3.25</td>
<td></td>
<td>0.599</td>
</tr>
</tbody>
</table>

Note. Statistics for global and local main effects are reported as well as data for the interactions. The mean, standard deviation, skew, cortiles, and Shapiro-Wilk test of normality p values are reported.
The distribution of data from each condition in Experiment 1 was tested for normality using a Shapiro-Wilk test. This test was selected because it was designed to test the normality of a distribution for researchers who are concerned about whether the skew of a distribution will affect statistical inferences. While each of the four distributions are skewed, the Shapiro-Wilk test of normality indicates that these distributions do not differ significantly from a normal distribution, all p > .05. For this reason, it is appropriate to make inferences on the data by using an ANOVA.
Experiment 2A and 2B

In user interface design, it is important to know more than which types of information users perceive first. In particular, user interface designers should know what information can help perform tasks more rapidly. Understanding what information users can utilize most rapidly can also shed light on what information is processed first. Which are more rapidly processed and utilized, global or local Gestalt principles of grouping? To speak to this question, Experiment 2 was conducted.

Method

Experiment 2 actually included two very similar experiments. In both experiments, unambiguous pushbutton targets were presented to participants for unlimited exposure durations. Common region and connectedness were each depicted as being relatively global in some conditions and local in other conditions. In Experiment 2A, I manipulated grouping scale (global versus local) common region. Examples of these stimuli are shown in Figure 28A and 28B. The common region in Figure 28A is twice as large in height as that in Figure 28B. Hence, Figure 28A represents global common region when compared to Figure 28B. In Experiment 2B, I manipulated grouping scale (global versus local) with connectedness. Examples of these stimuli are shown in Figure 28C and D. Connectedness in Figure 28C is four times as high than it is in Figure 28D. Figure 28C therefore represents global connectedness when compared to Figure 28D because it is larger in size scale.
Figure 28. Targets for Experiment 2

A

B

C

D

Note. (a) Global common region target for Experiment 2A. (b) Local common region target for Experiment 2A. (c) Global connectedness target for Experiment 2B. (d) Local connectedness target for Experiment 2B.

In each experiment, there were 12 trial blocks with 24 trials each. For Experiment 2A, as shown in Figure 29A and B, participants were shown an unambiguous global or local common region target until response. Response and reaction time were recorded and the target was followed by a 500ms mask screen. Figure 29A depicts the sequence in which the global common region target was shown. Conversely, Figure 29B depicts the sequence in which the local common region target was shown. For this experiment, participants were asked to “group the central pushbutton with the pushbutton that is located on the same region.” For Experiment 2B, as shown in Figure 29C and D, participants were shown an unambiguous global or local connectedness targets until response. Reaction time was recorded and the target was followed by a 500ms mask
screen. Figure 29C depicts the sequence in which the global connectedness target was shown. Conversely, Figure 29D depicts the sequence in which the local connectedness target was shown. For this experiment, participants were asked to, “group the central pushbutton with the pushbutton that is connected to it by bars.” Responses and reaction times were recorded.

Figure 29. Experimental sequences for Experiment 2

Note. (a) Global experimental sequence for Experiment 2A. (b) Local experimental sequence for Experiment 2A. (c) Global experimental sequence for Experiment 2B. (d) Local experimental sequence for Experiment 2B.
Results and discussion

Sixteen females volunteered to participate in Experiment 2A. In this experiment, I omitted the data from one participant because she did not accurately group over 10% of the unambiguous targets. The 15 remaining participants ranged in age from 18 to 31 with a mean age of 21.6. Nine females volunteered to participate in Experiment 2B. I omitted the data from one participant because she needed glasses to see the computer screen but did not have them with her during the experiment. The eight remaining participants ranged in age from 18 to 27 with a mean age of 21.3. Note that data from 15 participants was analyzed in Experiment 2A. After conducting this experiment and consulting with my dissertation advisor, I decided that 15 participants per experiment would not be necessary because of the large effect sizes. For this reason, Experiment 2B, which was conducted after Experiment 2A, included only eight participants.

For each experiment, reaction times to correct responses were compared by a within-subjects ANOVA. Responses were judged correct if the participant grouped the central pushbutton in the unambiguous stimulus by common region (Experiment 2A) or connectedness (Experiment 2B). Both experiments tested which type of grouping information was most readily processed and effectively utilized in speeded computerized tasks, as measured by reaction time. Figures 40A and 40B depict the results. In Experiment 2A, depicted in Figure 30A, grouping by global common region (M = 447 ms, SD = 83.4ms) was significantly faster than grouping by the local common region (M = 461 ms, SD = 82.4), F (1, 14) = 74.87, p < .001. In Experiment 2B, depicted in Figure 30B, grouping by global connectedness (M = 442 ms, SD = 114.0) was significantly

87
faster than grouping by the local connectedness (M = 449 ms, SD = 110.0), F (1, 7) = 6.35, p = .04. Taken together, the results of Experiments 2A and 2B suggest that global Gestalt grouping of both common region and connectedness is faster than local grouping.

Figure 30. Results for Experiment 2

Note. (a) Reaction time for global and local common region grouping. Results for Experiment 2A. Standard error indicated by bars. (b) Reaction time for global and local connectedness grouping. Results for Experiment 2B. Standard error indicated by bars.

Figure 31A depicts a box plot for global and local grouping reaction time means in Experiment 2A. The box plot provides additional information about the skew, quartiles and median of the distribution for each condition. These specific values can be found in Table 2. Additionally, Figure 31B and C depict distributions for all global and local reaction time scores respectively. These figures provide additional information about the distributions. As one can see, each of the two distributions is very similar because each has a positive skew. However, lower reaction time scores are more frequent for global grouping when compared to local grouping across prime durations. This explains why
global grouping was significantly faster than local grouping for the common region stimuli presented in Experiment 2A.

Figure 31. Additional analysis of Experiment 2A data

A

B

C

Global Grouping

Local Grouping

Note. (a) Box plot showing distribution of participant mean reaction times for global and local grouping conditions of Experiment 2A. (b) Frequency distribution for all Experiment 2A global reaction time trials.(c) Frequency distribution for all Experiment 2A local reaction time trials.
Table 2. Characteristics of RT distribution for Experiment 2A

<table>
<thead>
<tr>
<th>Experiment 2A</th>
<th>Mean</th>
<th>SD</th>
<th>Skew</th>
<th>5%</th>
<th>25%</th>
<th>50%/Med</th>
<th>75%</th>
<th>95%</th>
<th>S-W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>446.80</td>
<td>83.40</td>
<td>1.48</td>
<td>350.00</td>
<td>393.00</td>
<td>426.00</td>
<td>494.00</td>
<td>.</td>
<td>0.056</td>
</tr>
<tr>
<td>Local</td>
<td>461.07</td>
<td>82.43</td>
<td>1.43</td>
<td>360.00</td>
<td>408.00</td>
<td>445.00</td>
<td>508.00</td>
<td>.</td>
<td>0.063</td>
</tr>
</tbody>
</table>

Note. Statistics for global and local reaction time data are presented. The mean, standard deviation, skew, quartiles, and Shapiro-Wilk test of normality p values are reported.

The distribution of data from each condition in Experiment 2A was tested for normality using a Shapiro-Wilk test. While the distribution of data from each condition was positively skewed, the Shapiro-Wilk test of normality indicated that these distributions not differ significantly from a normal distribution, both p > .05. For this reason, it was appropriate to make inferences on the data by using an ANOVA.

Data from this experiment was further analyzed using a Wilcoxon Signed Ranks Test. This is a nonparametric test, which compares medians for correlated samples. The Wilcoxin Signed Ranks Test is considered to be a distribution free test which does not require the assumption of normality or homogeneity of variance. Further, this test is less influenced by the presence of outliers, when compared to an ANOVA. Since this was a repeated measures design, the two experimental conditions (global and local) could be considered to be correlated samples. Like the ANOVA used to compare the means, the Wilcoxon Signed Ranks Test revealed a significant difference between the global and local medians, Z = -3.41, p = .001. This means that according to this test, there was a significant difference between medians. Specifically, the global grouping reaction time median was significantly faster than the local grouping reaction time median as shown in Table 2.
Figure 32A depicts a box plot for global and local grouping reaction time means in Experiment 2B. The box plot provides additional information about the skew, quartiles and median of the distribution for each condition. These specific values can be found in Table 3. Additionally, Figure 32B and C depict distributions for all global and local reaction time scores respectively. Each of the two distributions is very similar because each has a positive skew. However, the global grouping distribution is shifted to the left, towards faster reaction times. This explains why global grouping was significantly faster than local grouping for the connectedness stimuli presented in Experiment 2B.

Figure 32. Additional analysis of Experiment 2B data
Note. (a) Box plot showing distribution of participant mean reaction times for global and local grouping conditions of Experiment 2B. (c) Frequency distribution for all Experiment 2B global reaction time trials.

Table 3. Characteristics of RT distribution for Experiment 2B

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Skew</th>
<th>5%</th>
<th>25%</th>
<th>50%/Med</th>
<th>75%</th>
<th>95%</th>
<th>S-W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>441.93</td>
<td>114.03</td>
<td>0.91</td>
<td>334.50</td>
<td>351.43</td>
<td>398.00</td>
<td>531.73</td>
<td></td>
<td>0.200</td>
</tr>
<tr>
<td>Local</td>
<td>448.96</td>
<td>109.98</td>
<td>0.92</td>
<td>351.80</td>
<td>357.15</td>
<td>408.15</td>
<td>531.45</td>
<td></td>
<td>0.136</td>
</tr>
</tbody>
</table>

Note. Statistics for global and local reaction time data are presented. The mean, standard deviation, skew, percentiles, and Shapiro-Wilk test of normality p values are reported.

The distribution of data from each condition in Experiment 2B was tested for normality using a Shapiro-Wilk test. For experiment 2B, the Shapiro-Wilk test of normality indicates that the two distributions, one from each condition, do not differ significantly from a normal distribution, both p > .05. For this reason, it was appropriate to make inferences on the data by using an ANOVA.

Data from this experiment was further analyzed using a Wilcoxon Signed Ranks Test. This was an important test to conduct for this particular experiment because the Shapiro-Wilkes test might not be accurate when sample size is small. In this experiment, the sample size was 8 participants per condition. Like the ANOVA used to compare the
means, the Wilcoxon Signed Ranks Test revealed a significant difference between the
global and local medians, $Z = -2.24$, $p = .03$. According to this test, the global grouping
median was significantly faster than local grouping median.
Experiment 3

It is important to note that a user’s grouping preferences can be influenced by a variety of factors. One such factor is prior grouping experience. Because Experiments 1 and 2 suggest that global grouping is more effective than local grouping, user interface designers may wish to incorporate global size scales into their displays. Will the incorporation of new size scales have utility for users who have had prior experiences with alternative size scales? Specifically, can prior grouping experiences with certain size scales influence how effective users are at grouping interface displays with other types of size scales? Experiment 3 was designed to answer this question.

Method

In Experiment 3, common region served as the global grouping principle and connectedness served as the local grouping principle (as in Experiment 1). As presented in Figure 33, participants were shown the ambiguous experimental primes and were trained to group by either global (common region) or local (connectedness). Primes were then followed by a mask screen that was shown for 500ms. One half of the participants were trained to group by the global scale properties of the ambiguous prime and one half were trained to group by the local properties. Training took place across six blocks of 24 trials. Immediately following the training period, participants were then asked to group the same experimental primes in a post-training experimental period. Half were asked to group by the grouping scale that they used during training and the other half were asked
to group by the alternative scale. The sequence depicted in Figure 33 remained the same for all practice and training conditions. Six blocks of 24 experimental trials were presented in each post-training period. Reaction times for the post-training period were measured. It was hypothesized that participants who practiced grouping by a particular scale would be faster to group by that scale in the post-training period when compared with those who were asked to group by an alternative scale.

Figure 33. Experimental sequence for Experiment 3

\[ \text{Experimental Prime is followed by a mask.} \]

Results and discussion

Sixty-one participants volunteered to participate in Experiment 3. I omitted the data from one participant from the global to local condition because her reaction times were more than three standard deviations above the group mean. The 60 remaining participants ranged in age from 18 to 45 with a mean of 20.6. Four males and 56 females participated. A 2X2 between participants ANOVA (Training size scale X Test size scale)
was used to compare reaction times of participants who used an alternative size scale after training with reaction times of participants who used the same size scale. There was no main effect for practice condition (global or local), $F < 1$, or test condition (global or local), $F < 1$. Furthermore, there was no significant interaction, between the practice and test conditions, $F < 1$. These results suggest that participants in this experiment performed equally well at grouping by a particular scale (global or local) in a test phase when trained by the same scale or an alternative scale in a preceding practice session.

Figure 34 depicts a box plot for all conditions in Experiment 3. The box plot provides additional information about the skew, quartiles and median of the distribution for each condition. These specific values can be found in Table 4.

Figure 34. Box plot showing distribution of mean reaction times for all conditions of Experiment 3. Condition ‘GG’ was one in which a global practice session was followed by a global test session. Condition ‘LG’ was one in which a local practice session was followed by a global test session. Condition ‘LL’ was one in which a local practice session was followed by a local test session. Condition ‘GL’ was one in which a global practice session was followed by a local test session.
Table 4. Characteristics of RT distribution for Experiment 3

<table>
<thead>
<tr>
<th>Experiment 3</th>
<th>Mean</th>
<th>SD</th>
<th>Skew</th>
<th>5%</th>
<th>25%</th>
<th>50%/Med</th>
<th>75%</th>
<th>95%</th>
<th>S-W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glo Practice Main</td>
<td>417.37</td>
<td>48.80</td>
<td>1.57</td>
<td>356.40</td>
<td>383.75</td>
<td>415.50</td>
<td>440.75</td>
<td>535.45</td>
<td>0.010</td>
</tr>
<tr>
<td>Loc Practice Main</td>
<td>429.37</td>
<td>46.67</td>
<td>0.45</td>
<td>356.40</td>
<td>398.00</td>
<td>415.50</td>
<td>468.00</td>
<td>509.45</td>
<td>0.040</td>
</tr>
<tr>
<td>Glo Test Main</td>
<td>418.93</td>
<td>49.95</td>
<td>1.70</td>
<td>352.00</td>
<td>389.00</td>
<td>404.50</td>
<td>429.50</td>
<td>545.90</td>
<td>0.010</td>
</tr>
<tr>
<td>Loc Test Main</td>
<td>427.80</td>
<td>45.81</td>
<td>0.19</td>
<td>361.10</td>
<td>383.00</td>
<td>426.00</td>
<td>465.75</td>
<td>509.45</td>
<td>0.231</td>
</tr>
<tr>
<td>Glo Pra x Glo Test</td>
<td>413.80</td>
<td>55.93</td>
<td>2.42</td>
<td>352.00</td>
<td>385.00</td>
<td>398.00</td>
<td>428.00</td>
<td></td>
<td>0.010</td>
</tr>
<tr>
<td>Loc Pra x Glo Test</td>
<td>424.07</td>
<td>44.54</td>
<td>0.79</td>
<td>352.00</td>
<td>398.00</td>
<td>406.00</td>
<td>455.00</td>
<td></td>
<td>0.062</td>
</tr>
<tr>
<td>Loc Pra x Loc Test</td>
<td>434.67</td>
<td>49.68</td>
<td>0.19</td>
<td>360.00</td>
<td>398.00</td>
<td>434.00</td>
<td>480.00</td>
<td></td>
<td>0.489</td>
</tr>
<tr>
<td>Glo Pra x Loc Test</td>
<td>420.93</td>
<td>42.17</td>
<td>0.01</td>
<td>362.00</td>
<td>376.00</td>
<td>425.00</td>
<td>465.00</td>
<td></td>
<td>0.396</td>
</tr>
</tbody>
</table>

*Note.* Statistics for all main effects and interactions are presented. The mean, standard deviation, skew, quartiles, and Shapiro-Wilk test of normality p values are reported.

The distribution of data from each condition in Experiment 3 was tested for normality using a Shapiro-Wilk test. This test suggests that the data distributions for the practice factor, one from each condition, differ significantly from a normal distribution [Global (30) = .89, p = .01, Local (30) = .92, p = .04]. Furthermore, the Shapiro-Wilk test of normality indicates that the data distributions for the test factor differ significantly from a normal distribution for one condition but not for the other [Global (30) = .85, p = .01, Local (30) = .95, p = .23]. Since one of the assumptions of an ANOVA is that the data in each condition is normally distributed, inferences on the data could not be made using a traditional ANOVA. For this reason, a transformation of the data was needed to reduce the skew. This transformation would maintain the relationships between all the raw data, but correct for severe skew.

Initially, the data were transformed by use of square root. This transformation was insufficient because it did not bring all distributions to normality according to the Shapiro-Wilk test. Next, the data were transformed logarithmically. Again, the Shapiro-Wilk test revealed that all distributions had not reached normality. Finally, a negative
The reciprocal transformation of the data was used. This transformation of the data changed the distributions so they no longer differed significantly from the normal distribution according the Shapiro-Wilk test. Note that this transformation is one of the most powerful but it was necessary to achieve a normal distribution for each condition. The transformed data were then analyzed using an ANOVA. There was no main effect for practice condition (global or local), $F(1, 56) = 1.11, p = .30$, or test condition (global or local), $F < 1$. Furthermore, there was no significant interaction, between the practice and test conditions, $F < 1$. Consistent with the ANOVA conducted before the data transformation, these results suggest that participants in Experiment 3 preformed equally well at grouping by a particular scale (global or local) in a test phase when trained by the same scale or an alternative scale in a preceding practice session.
General Discussion

Research that has examined the effect of the Gestalt grouping principles on display design has largely influenced how designers group elements in user interface displays. However, limitations set by task demands have been shown to constrain the amount of time that users can preview interface displays. Hurrying to meet deadlines, quick saccades between displays to integrate information, unwanted information sorting and repetition of familiar tasks are all examples of limitations on user interface viewing time. Recent discoveries in vision science, suggest that grouping can occur in more than one way across time (Gulick & Stake, 1957; Kurylo, 1997; Moore & Brown, 2001; Navon, 1981; Oliva & Schyns, 1997; Rauschenberger & Yantis, 2001; Sanocki, 1993; Schulz, 2002; Schulz et al., 2001; Schulz & Sanocki, 2003; Schyns & Oliva, 1994). These discoveries that address exposure time limitations and grouping, have not been taken into account in user interface research. For this reason, in the reported experiments, time limitations were carefully considered with respect to the grouping of elements presented in user interface displays.

Experiment 1 was designed to determine whether grouping operates differentially on global and local information as a function of exposure duration. Observers were presented with ambiguous prime displays, in which a central pushbutton could group with a pushbutton on one side by a global grouping property (common region) or with a pushbutton on the opposing side by a local grouping property (connectedness). The
primes were masked to limit prime duration and then followed by an unambiguous target display in which the central pushbutton could group with the pushbutton to the right or to the left. Participants reported whether the central pushbutton of the target display grouped with the right or left pushbutton. Targets matched either the global or local grouping interpretation of the preceding prime.

I used difference scores to examine the influence of the primes. I calculated the mean global grouping advantage for each prime (see Figure 22 for primes) by subtracting the mean RT for experimental global match targets from the mean RT for global control targets (see Figure 23A for global targets), for each prime duration. Likewise, I calculated the mean local advantage for each prime by subtracting the mean RT for experimental local match targets from the mean RT for local control targets (see Figure 23B for local targets), for each prime duration. Positive scores signified an advantage whereas negative scores signify a disadvantage.

The results of Experiment 1 revealed that there was a main effect of size scale, global versus local. This result suggested that participants perceived global groupings to be more salient than local groupings in both prime durations. There was no significant main effect of prime duration 50ms versus 4000ms on reaction time. There was no significant interaction of grouping scale and prime duration. Taken together, the results 50ms and 4000ms prime durations are consistent with the notion that global groupings are more salient than local groupings across prime durations. This suggests an overall global grouping advantage.
Experiments 2A and 2B were designed to determine whether global or local grouping was faster in unlimited exposure conditions. Participants were shown unambiguous targets for an unlimited amount of time. The targets contained global or local grouping scales with either common region (Experiment 2A) or connectedness (Experiment 2B). Participants were asked to simply group each unambiguous target. Reaction times for correct grouping responses were analyzed. The results of Experiment 2 suggest that global Gestalt grouping of both common region and connectedness is much faster than local grouping.

Taken together, the results of Experiments 1 and 2 have some interesting implications. Knowing which grouping principles are most effectively seen and utilized can help designers determine which grouping principles to present in interfaces that are designed for fast searches. To expand on this idea, let us reconsider the example of the speeded database file search task that was mentioned earlier in this paper. Recall that the employee in our previously mentioned example was searching for a target ‘Picture’ file in the “Presentation” branch of a database (Figure 8). The employees’ search for the target file was speeded because of a deadline that he attempted to meet. However, the target file was ambiguous as to how it could be grouped within the directory branches. Specifically, the target ‘Picture’ file could be grouped by connectedness with the ‘Presentation’ branch of the directory or by common region with the ‘Personal’ branch of the directory. Due to this ambiguity, the employee would likely need to slow search to make certain that the correct file was retrieved from the correct branch of the database.

101
How can the needs of the employee in this example be better addressed? The employees’ task of meeting a deadline requires him to search for the target file in the database very quickly. Can the design of this database be altered to accommodate fast searches? The reported research can be used to address these types of questions. Since global grouping was shown to be salient across exposures and most usable for a speeded task, global grouping could be incorporated into the design of interface displays that are used and previewed in speeded task conditions. The employee database that was initially depicted in Figure 8 has been redesigned in Figure 35 to facilitate grouping in fast file searches. As one can see, global common region is used. Based on the reported results, employees who are attempting to perform speeded searches for files within the directory branches will likely perceive and most efficiently rely on the global (common region) grouping.

Figure 35. Improved database

Note. Redesign of database presented in Figure 8.

In Experiment 3, prior experience with a particular size scale (global or local) was compared to performance when using an alternative size scale. Experiment 3 was designed to determine whether users have an advantage for grouping by the scale that they are accustomed to using and whether there is a disadvantage when designers change...
the size scale within an existing interface display. Participants in this experiment were shown the ambiguous experimental primes and trained to group by a particular Gestalt size scale (global or local). One half of participants were trained to group by global properties and one half of participants were trained to group by local properties. A test phase then followed the practice session and reaction times were recorded. Participants in this experiment preformed equally well at grouping by a particular scale (global or local) in a test phase when trained by the same scale or an alternative scale in a preceding practice session.

It is interesting to think about why null results were found in Experiment 3. One of the most likely reasons is that participants may not have had enough practice with each particular size scale for this practice to have had an effect during the test phase. In this experiment, participants only had about 10 minutes of practice with a particular size scale before they were tested. This may not have been enough time to become accustomed to using a particular size scale. In the real world, in situations in which users have years of practice using a particular size scale, changing the size scale may dramatically reduce efficiency for using new displays. Perhaps if participants had been trained to group by a particular scale over an extended period of time, during multiple practice sessions, the grouping results at test for alternative scales would have been significantly slowed.

Contribution to engineering

The reported results contribute to a greater understanding of how grouping should be used in user interface displays. According to the three pilot studies, reliance on global
and local grouping can change across time. Specifically, global grouping was predominantly chosen in the shorter stimulus durations and a mixture of global and local groupings were chosen in the longer exposure durations. These results suggest that grouping may not be the static display property that designers once might have thought. Rather, they suggest that grouping is quite a bit more complex and that perhaps time limited task demands should be considered before display groupings are selected. According to the results of the three pilot studies, when task demands limit exposure to user interface displays, global groupings should be implemented.

The results of dissertation Experiment 1 reveal that global grouping can be more salient than local grouping across prime duration. These results suggest that global Gestalt groupings are should be incorporated into computerized displays because they are more easily seen. The results of dissertation Experiment 2 reveal that when specific grouping principles are depicted more globally, users can interact with displays more quickly. These results also suggest that global grouping should be incorporated into user interface displays to speed interaction.

One might notice that the advantages found for global grouping facilitation in Experiment 1 and speed in Experiment 2 were not very large and varied only by milliseconds. Based on this small difference, the incorporation of global grouping in interface displays might not appear to make a dramatic difference on the usability. When considering these ideas, it is important to not only think of the application of these results in terms of the individual user. Instead, it is important to think of how these results could apply to a user population. For example, the web-based email communication program
“Hotmail” has been reported to have over 30 million users. Saving these users even a few milliseconds per mouse click could amount to a huge time savings overall. It should be made clear that the application of global groupings will be a small step toward the larger goal of creating a more seamless interaction between humans and machines. This relatively small advance will be combined with other small advances in the area of grouping and user interface to change the speed of progress that humans can make with machines.

It is important to note that the reported results should not be viewed as the definitive organizational solution for every computerized display. That is to say, these results should not be used to redesign currently functional displays by taking away existing attributes that might make them easy to use. However, these results can be applied within the context of each specific display to facilitate user interaction. For example, if a factory display is designed to have a good match between system and the real world, it would not be wise to strip that display of this positive attribute and implement global Gestalt grouping in place. Rather, perhaps important elemental groupings within this display could be depicted as more global. In this way, the good match between system and real world would be persevered.

The reported pilot results may call for modifications to modern ideas about the application of grouping in user interface design. For example, the Proximity Compatibility Principle (PCP) would likely need to be amended to take processing time into account. The reported results could fit nicely into the PCP if processing proximity would call for closeness in the time sensitive form of perceptual proximity. That is to say,
perceptual proximity would be determined by how elements are perceived as being
globally grouped depending on the time sensitive task demand. With reference to the first
pilot study for example, information that is close in processing proximity and presented
in limited exposure durations would be best grouped by perceptual closeness of color.
This is because color was perceived to be perceptually proximal at short stimulus
durations. Conversely, information that is close in processing proximity and presented in
longer exposure durations could be best grouped by color or connectedness. This is
because both color and connectedness were perceived to be perceptually proximal at
longer stimulus durations.

The reported results also fit well within the context of user interface research.
When researchers in user interface try to compare Gestalt principles with one another, to
determine salience or predominance of particular grouping principles, it may be really
that they are comparing global and local interpretations of these principles. For example,
Addy (2000) found that in unlimited exposures, participants best grouped by color
similarity, followed by common region, then followed by proximity. Addy (2000) had
shown that, in certain conditions, some grouping principles could be stronger than others.
Perhaps in these experiments, color similarity was perceived as the most global grouping
principle, followed by common region and proximity. If so, Addy’s findings could be
explained by the reported results.

As one can see, the reported results do make a contribution to the field of
engineering. However, the reported results are not just interesting within this context.
They are interesting within the context of theoretical Gestalt grouping research in vision science as well.

**Contribution to vision science**

Researchers in human vision science have been interested in the study of global and local processing of elements in spatial scenes. They have studied the relative contribution of large and small scale shapes to the eventual development of a percept. These researchers have discovered that global scene details are processed more rapidly than local scene details (Navon, 1981; Sanocki, 1993). Recall that Navon (1981) showed that at short stimulus durations, it was easier for participants to identify the global (rather than local) features of a perceptual stimulus when compared to longer stimulus durations. Recall also that, Sanocki (1993) found that global information contributed more to object identification earlier in processing and much less so in later processing. The results of the reported pilot experiments 1 to 3 add to the previous findings by suggesting that the previously discovered time course trend may not be limited to the processing of general global and local scene details. Rather, they suggest that this trend may also extend to the perceptual grouping of global and local scene properties.

The results of the pilot studies 1 to 3 also suggest that yet another grouping attribute may be processed along a time course. Here, the attribute is the size scale (global or local) of grouping. This finding fits nicely within the time course literature that has found grouping can change relative to size (Gulick & Stake, 1957), low versus high frequency information (Schyns & Oliva, 1994, Oliva & Schyns, 1997), global versus
local configuration (Navon, 1981; Sanocki, 1993), object completion (Rauschenberger &
Yantis, 2001), lightness and transparent filters (Moore & Brown, 2001), grouping by
proximity and alignment chromatic color, achromatic color and shape similarity (Kurylo,
1997; Schulz, 2001; Schulz, 2002; Schulz et al., 2001; Schulz & Sanocki, 2003).

The results of Experiment 1 suggest that global information can be more salient to
people than local information across durations. The results of Experiment 2 suggested
that people can be faster at interacting with global information. If these results could
generalize to real world displays (as opposed to simple computer screen layouts), they
might suggest that humans find global groupings to be more salient and usable when
compared to local grouping in real world spatial environments. Evolutionarily speaking,
this type of finding would be of interest. Perhaps large scale shapes have been more
evolutionarily important to humans than fine scene details. For example, it would
probably be more important for a caveman to determine whether a large object in the
distance was an animal when compared to making an assessment about the sharpness of
its teeth.

Future directions

The results observed in the reported experiments raise several interesting
questions for future research. First, it would be interesting to test global and local
groupings of Gestalt principles other than common region, connectedness and color
similarity. This would help determine whether global salience can generalize to other
grouping principles. For example, global and local grouping by proximity and other
forms of similarity could also be tested using the reported methods. If the reported results could be shown to generalize to other Gestalt principles, perhaps interface designers could plan to globalize elemental groupings as a general rule for enhanced display design.

Another interesting research avenue would be to test global and local groupings in displays that more closely resemble user interface. While the displays presented in the reported experiments contained pushbuttons, they were much more basic than a typical user interface display. In contrast, a typical interface display would contain a variety of elements, in addition to pushbuttons, like links, text and pictures. Testing global and local groupings in displays that more closely resemble user interface could make for less of a basic research program but may generate information that is more suitable for actual user interface displays. In this more applied research, alternative types of measures could be used. These measures could include navigation speed for completing particular tasks, measurement of eye fixation and saccades between display elements, and visual search speed for the location of target pushbuttons. Each one of these measures could provide interesting new information about human performance with respect to global and local interface groupings.

As an expansion on the previous idea, it would be interesting to determine how well the reported results would generalize to real world physical displays. There are many types of physical displays in real world spatial environments including factories, airplanes and control towers. Similar to designers of user interface, designers of these types of displays incorporate elemental groupings to enhance user performance. Perhaps
human interactions with these types of displays could be enhanced by an understanding of human perception of global and local Gestalt groupings.

Finally, it is important to note that the reported experiments have served to detail the effect of one task factor, processing time, on global and local perceptual grouping. However, there may very well be influences of other task factors, like attention and distractions. These task factors should also be carefully considered within the context of user interface displays.

Overall summary

Across a series of experiments, the nature of global and local perceptual grouping was examined within the context of user interface displays. Three pilot studies revealed that global grouping was more salient in short stimulus durations when compared to longer stimulus durations. The reported dissertation experiments showed that global grouping was more salient overall and more usable for speeded tasks.

The reported results contribute to our understanding of how to better design user interface displays. The application these findings to user interface design may enhance our ability to perceive and comprehend information presented on computerized displays. This application may improve our accuracy in working with computerized systems, therefore increasing the speed in which we perform computer related tasks. Further, the application these results may increase productivity in many types of work environments that rely on computers and ease interactions with computerized devices used for personal activities, such as cell phones and PDAs. The reported results also are important within
the context of vision science research. They further our understanding of global and local scene processing by allowing us to consider this processing type of processing with respect to grouping.

Perceptual grouping may be malleable and moderated by a host of task factors. It is important to carefully study the effect of the many task factors that moderate perceptual grouping. The discoveries that we make in this new research domain will likely enhance how humans perceive and interact with information in computerized and real world environments.
References


*Psychological Review, 96 (3),* 433-458.


Schulz, M. F. (2001). *Time course of perceptual grouping: Lightness constancy and amodal completion*. Interdepartmental Honors Thesis. Published in the University of South Florida Library.


*Psychological Science, 14*(1), 26-30.

Schulz, M. F. *Time course of perceptual grouping: A priming study.* (2002)


(Original work published 1923).


About the Author

Melissa F. Schulz received a Bachelor’s Degree in Psychology from the University of South Florida in 2001 and a M.A. in Cognitive Psychology in 2002. She started teaching Research Methods Laboratory and continued on as an instructor of Sensation and Perception at the University of South Florida. She graduated from the Ph.D. program at the University of South Florida in 2004.

While in the Ph.D. program at the University of South Florida, Dr. Schulz was very active in helping her students reach their highest potential. She also first authored a publication in *Psychological Science* and made several talks and poster presentations at international Vision Science and Teaching of Psychology conferences.