THE ANALYSIS OF SLANT-FROM-TEXTURE IN EARLY VISION

by

Deborah J. Aks
University of British Columbia

B.A., The State University Center of New York at Binghamton, 1984
M.A., The University of British Columbia, 1988

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in
THE FACULTY OF GRADUATE STUDIES
(Department of Psychology)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
August 1993
© Deborah J. Aks
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.
Abstract

A considerable amount of research exists on the subjective perception of three-dimensional structure from texture gradients. The present set of experiments extends these tests of phenomenal perception by examining the underlying processes used in interpreting slant-from-texture. The first two experiments show that measures of subjective perception predict speeded performance in a visual search task, and that the mediating representation relies on an assumption of projective size (i.e., discriminating the size of the target is difficult when the short target is far or the long target is near). The third experiment shows that sensitivity to apparent depth in the texture display is present even in rapid and parallel search conditions where early vision is known to operate. The fourth experiment assesses the relative contribution of two dominant dimensions of the texture gradient -- "perspective" (i.e., a radial pattern) and "compression" (i.e., a foreshortened pattern). Both dimensions are detected by early vision as signals for apparent depth. The final experiment examines how early vision codes these two dimensions. Sternberg's (1969) Additive Factors Method (AFM) is used to assess separability of encoding, and Blalock's path analysis (1962, 1985) is used to examine the order of encoding. AFM shows that perspective and compression have independent influences on search performance in the most rapid search conditions, but that their interaction increases as search slows. The path analysis shows further that when both texture dimensions are available, perspective exerts a more immediate and perhaps even an exclusive influence on performance. These findings support the view that perspective and compression are coded separately at the earliest stages of visual processing and share a common code only later in visual processing.
Table of Contents

Title Page.............................................................................................................. i
Abstract ................................................................................................................... ii
Tables of Contents ................................................................................................ iii
List of Tables ......................................................................................................... vii
List of Figures ....................................................................................................... viii
Acknowledgments ................................................................................................. x
Chapter 1
   Slant-from-Texture ......................................................................................... 1
   Overview ............................................................................................................ 4
   Texture gradients ............................................................................................. 4
   Perceptual use of texture gradient information ................................................ 7
       Empirical tests. ................................................................................................ 8
Chapter 2
   Separating the process from the product ......................................................... 11
   Do subjective measures predict speeded performance? .................................... 14
   The visual search task ..................................................................................... 15
   Experiment 1
   Slant-from-texture influences on visual search ............................................... 17
   Method .............................................................................................................. 18
       Subjects ......................................................................................................... 18
       Stimuli and Procedure ................................................................................. 18
   Results .............................................................................................................. 21
       Display size effects........................................................................................ 25
       Central location effects ............................................................................... 25
       Location and testing order .......................................................................... 26
   Discussion ......................................................................................................... 26
   Experiment 2
   Separating top-to-bottom from apparent-distance effects ............................ 27
   Method .............................................................................................................. 28
       Subjects ......................................................................................................... 28
       Stimuli and Procedure ................................................................................. 28
Slant - from - texture
iv

Results .................................................................................................................. 28
  Location effects in the control condition .......................................................... 32
  Display size effects ............................................................................................ 32
  Testing order ...................................................................................................... 33
Discussion ........................................................................................................... 33

Chapter 3
How early are the slant-from-texture influences? ........................................... 35
A two stage model of vision .............................................................................. 35
Encoding slant-from-texture
  Early or late? ..................................................................................................... 37
RT Slope performance ....................................................................................... 38
  Projective size or apparent proximity? ........................................................... 38
Summary ............................................................................................................ 41

Experiment 3
Slant-from-texture influences in early vision ................................................... 41
Method ................................................................................................................. 42
  Subjects ............................................................................................................ 42
  Stimuli and Procedure ..................................................................................... 42
Results ................................................................................................................. 46
  Threshold for preattentive vision .................................................................... 46
  Visual search on textured surfaces ................................................................. 46
  Display size and apparent depth ................................................................... 48
Additional influences .......................................................................................... 51
  Practice ............................................................................................................. 51
  Top - bottom search biases ......................................................................... 51
  Left - right search bias .................................................................................. 53
  Item Orientation ............................................................................................... 54
  Size - depth consistency ................................................................................ 56
Discussion ............................................................................................................ 57
  Item grouping? ............................................................................................... 57
  Search biases ................................................................................................. 58
  Implications ..................................................................................................... 58
Chapter 4

An examination of the texture gradient dimensions

Monocular Depth Cues
Decomposing the texture gradient
Relative contribution to subjective perception

Experiment 4
Influences from dimensions
Method
Results

Perspective & Compression Combined
Item orientation

Perspective Results
Display size and apparent depth
Performance difference across backgrounds
Item orientation
Additional location influences
Summary of perspective effects

Compression Results
Performance difference across backgrounds
Item orientation
Additional influences
Summary of compression effects

Discussion

Chapter 5
How does our visual system combine the dimensions?
A test of additivity
Method
Results

Contrast test of Additivity
Processing speed and additivity
Item orientation
Speed of Processing, additivity, and item orientation

Discussion

Processing sequence
Method ................................................................. 104
Results ................................................................. 105
Discussion ............................................................. 109
General Discussion .................................................. 111
  Implications for theories of early vision ..................... 112
  Implications for theories of surface perception ............ 114
  Early vision and surface perception ......................... 116
References .................................................................. 118
Appendix A .............................................................. 129
List of Tables

Table 1: Standard Error of the Mean (SEM) for RT and percentage of errors in Experiment 1 ................................................................. 23

Table 2: Standard Error of the Mean (SEM) for RT and percentage of errors in Experiment 2 ................................................................. 31

Table 3: Mean correct reaction time and percentage of errors in Experiment 3 -- Combined Gradients ................................................................. 47

Table 4: Mean correct reaction time and percentage of errors in Experiment 3 -- Top-Bottom Analysis ................................................................. 52

Table 5: Mean correct reaction time and percentage of errors in Experiment 4 -- Perspective Gradient ................................................................. 76

Table 6: Mean correct reaction time and percentage of errors in Experiment 4 -- Compression Gradient ................................................................. 80

Table 7: Inter-correlation and partial correlations across texture gradient conditions ................................................................. 106
List of Figures

Figure 1: Picture containing slant-from-texture. .................................................. 2
Figure 2: Construction of a texture gradient. ......................................................... 3
Figure 3: Trigonometric function relating surface texture to slant angle. ............... 5
Figure 4: Visual search displays. ........................................................................... 20
Figure 5: Experiment 1 results -- Mean correct RT and Percent Errors. ................. 22
Figure 6: Experiment 2 results -- Mean correct RT and Percent Errors. ................. 30
Figure 7: Experiment 2 RT Slope performance. ..................................................... 39
Figure 8: Combined texture gradient background used in Experiment 3. ............... 43
Figure 9: Full set of search items used to assess thresholds for early vision. .......... 44
Figure 10: Experiment 3 results -- Mean correct RT and Percent Errors. .............. 49
Figure 11: Experiment 3 results separated across item orientation. ...................... 55
Figure 12: Perspective, compression and combined gradient search displays.
    Figure 12a: Vertical items. ................................................................. 63
    Figure 12b: Horizontal items. ............................................................. 64
Figure 13: Mean RT difference and Percent Error difference across slant and control
    background for Perspective, Compression and Combined gradient target
    trials in Experiment 4.
    Figure 13a: Short target trials. ............................................................. 70
    Figure 13b: Long target trials. ............................................................ 71
Figure 14: Mean RT and percent error difference separated across item orientation.
    Figure 14a: Short target trials.............................................................. 73
    Figure 14b: Long target trials. ............................................................ 74
Figure 15: RT slopes for Perspective, Compression and Combined conditions. ....... 83
Figure 16: Additivity test using orthogonal variation of texture dimensions separated across size-depth consistency.

Figure 16a: Consistent conditions......................................................... 92
Figure 16b: Inconsistent conditions....................................................... 93

Figure 17: Contrast scores -- Mean difference in RT and percent error across all texture (perspective and compression) present and absent trials.............. 95

Figure 18: Additivity results separated across item orientation.

Figure 18a: Vertical items................................................................. 97
Figure 18b: Horizontal items......................................................... 98

Figure 19: Contrast scores separated across item orientation.

Figure 19a: Short target trials......................................................... 100
Figure 19b: Long target trials...................................................... 101

Figure 20: Path representation of texture processing.............................. 107
Acknowledgments

I extend great thanks to all of the following people who helped me through quite a journey -- Jim for teaching new and inspiring approaches to old perceptual problems and for the tremendous contributions to the Slant-from-texture research; Stan for providing creative ideas, a wealth of illusions and endless challenges throughout my graduate career; Janet, Rob and Rod for keen insight on important issues in my thesis and much appreciated encouragement; Dimitri for superb editorial contributions; Roger for trigonometric expertise; Chris for Reason from around the globe; my family for endless emotional support -- Mom for reminding me of the applied, the virtual and the real worlds; Dad for reminding me of the sailing and scuba worlds; my sibs -- Judy, Steve and Lori -- for being my friends; my friends -- Jill, Kelly, David, Sarah, Ed, Shuji and all the folks from the Vision lab & the Big Apple -- for being so supportive. And last but not least... Mohamad for his omnipresence.
Chapter 1: Slant-from-Texture

When viewing a picture of a textured surface, such as the snow-rippled mountain in Figure 1, observers typically report an impression of a slanted surface receding into the distance. This is true regardless of the fact that a two-dimensional (2-D) image can never completely specify a surface in depth. How do we perceive depth from an image that is constrained to two dimensions? One potential solution is that our visual system uses geometric rules to interpret 2-D images as projections from real objects in the world (i.e., Brunelleschi; 1413, cited in Kemp, 1978). Such a strategy would imply that our visual system assumes a single vantage point, and uses rules of linear perspective when interpreting line drawings. This may mean that vision interprets the vanishing points in a 2-D image as distance points in the 3-D world and uses this information to determine geometric relationships of objects in space. A detailed description of ways in which we may solve this correspondence problem between the image and the scene is presented following a brief discussion of the projective distortion in the texture gradient and an overview of the present studies.

When light from a slanted surface is projected onto an image plane, uniform surface texture is distorted in the image to form a texture gradient as shown in Figure 2. In a flat textured surface, the horizontal and vertical dimensions of texture elements diminish gradually in size as the slant angle increases away from the observer. The systematic variation of texture due to changes in distance and slant provides important information about the 3-D properties of an object relative to an observer. The relative distance information specified by texture gradients will be referred to as slant-from-texture throughout the thesis. There is also information about the magnitude and direction that a surface dips away from the frontal plane but these are of secondary importance in this study of apparent depth perception.
Figure 1. Picture of mountain glacier as an example of a textured surface receding into the distance. Horizontal and vertical distortions of the snow ripples in the image indicate information about surface slant.
Figure 2. Representation of light from a flat - slanted surface being projected onto an image plane showing how uniform surface texture is distorted in the image to form a texture gradient. The horizontal and vertical dimensions of texture elements in the image diminish gradually in size as the slant angle of the floor increases away from the viewer.
Overview

The perception of slant-from-texture was investigated in a series of visual search experiments with the goal of answering the following questions. First, do tests of subjective perception and rules of projective geometry predict visual search performance in slant-from-texture displays? Second, which visual systems are influenced by slant-from-texture: early or late? Third, which dimensions of the texture gradient influence visual search, and does the literature on subjective perception predict the relative importance of these dimensions in search performance? Fourth, which dimensions can be detected in parallel in visual search? The final section examines how the visual system combines the dimensions of the texture gradient to elicit an impression of slant. Implications for underlying representations and processes are discussed.

Texture gradients

A considerable amount of work has been devoted to mathematical and geometrical proofs relating surface slant of objects to texture gradients in images (i.e., Cutting and Millard, 1984; Dunn, 1986; Flock, 1964; 1965; Kantani, 1984; Marr, 1982; Pizlo & Rosenfeld; 1992; Purdy, 1960; Stevens, 1981; 1983a; 1983b; Witkin, 1981). Cutting and Millard's (1984) trigonometric functions representing the perspective and compression aspects of texture are shown in Figure 3. To use these equations in extracting slant-from-texture, assumptions about texture regularity, texture element spacing, and surface planarity are required (e.g., texture elements are assumed to be regular in size, uniform in spacing, and flush against a flat surface). Given these constraints, corresponding rules can be used to describe how the visual system interprets texture gradients as having slant and distance information.

J.J. Gibson (1950a; 1950b; 1979) is credited as the first researcher to test whether texture is mathematically and psychologically sufficient as a stimulus for depth
Figure 3. Cutting and Millard's (1984) trigonometric functions representing the perspective and compression aspects of the surface texture and their relationship to surface slant.
perception. A computational analysis by Purdy (1960) demonstrated that four variations among optical texture elements were equally informative: gradients of size, compression, convergence, and density. In each case, the slant of a planar surface is specified by the \( \text{arccot} \left( \frac{G}{k} \right) \) where \( G \) is the gradient and \( k \) is a constant specific to the type of gradient. Purdy's equations, although they accurately describe gradient information, do not separate out additional information present in the gradient that may more accurately be used by the human visual system to judge the slant of surfaces.

Recognizing the insufficiency of these gradient solutions, Stevens (1981) argues instead that the visual system uses separate dimensions of surface texture to estimate distance (element width), and orientation (element height / width). This approach eliminates the confound of distance and orientation information, which when combined can lead to erroneous judgments of slant. More precise estimates of surface slant are obtained by separating orientation into tilt and slant components, and mapping them directly onto object orientation and distance. In Stevens' framework, tilt (i.e., the direction in which the surface recedes) is easy to determine from a texture gradient, but slant (i.e., the amount which the surface recedes) is difficult to find. Tilt is reliably determined by first detecting the direction with the least variation in texture spacing -- this is the line that passes through equal image heights of the texture elements, and is equidistant from the viewer. The direction perpendicular to this line is the surface tilt.

The rules for determining slant, however, change dramatically depending upon the properties of the surface texture. For a flat surface with regular texture elements (i.e., tiled floor), \( \text{Slant} = \text{arccos} \left( \frac{\text{element height}}{\text{element width}} \right) \). For a surface with protruding elements (i.e., grass), \( \text{Slant} = \text{arcsin} \left( \frac{\text{element height}}{\text{element width}} \right) \). Stevens's work explicitly shows that the calculation of slant, unlike tilt, is unreliable across a variety of surfaces. He shows that the number of possible mathematical
functions increases with each additional exception to the "flat and uniformly textured standards".

Stevens (1983) suggests a rule based on the initial determination of tilt. This rule relies on use of a texture dimension that is equidistant from the viewer and which is at right angles to the axis of tilt (See tilt calculation). Stevens refers to this dimension as the characteristic dimension (CD), and as shown in Chapter 4, it eliminates the need for an assumption of texture elements resting flat on a surface. Thus, the CD is simply a function of the distance from the viewer and, as demonstrated later, is equivalent to a primary dimension of the texture gradient (i.e., perspective). From this source of information, slant can be calculated by determining the gradient (G) of CD’s using the equation: \( \text{Slant} = \arctan \left( \frac{G \times \text{CD}}{\text{CD}} \right) \).

The simplest and most general description of the reliable relationship between texture size and distance is based on the fundamental property of projective geometry that "far objects project smaller images on the retina than close objects of the same size". Projective size can also be described as the relative changes in the visual angle of image elements, defined as the \( \arctan \left( \frac{h}{d} \right) \) where \( h \) is the height of the focal stimulus and \( d \) is the distance from the observer to the stimulus. This relationship, however, has at least one limitation -- its imprecision. Unlike the other equations which provide absolute slant information, projective size only provides relative distance information. Reliance solely on principles of projective geometry, a somewhat gross measure of distance, is plausible in light of evidence (some of which is summarized below) that humans are better at estimating relative rather than absolute distance.

**Perceptual use of texture gradient information**

Having established the sufficiency of texture gradients to specify distance and slant information (given the noted assumptions), our next concern is whether or not perceivers use this information. This question is far from new. This controversial issue
has been debated among perceptual researchers since the middle of the nineteenth century when Helmholtz argued that some laws of physical optics are manifest in unconscious processes of shape constancy (1868/1968). Similar claims include Gibson's belief in the direct perception of projective invariants (1950, 1979), Rock's view that the visual system is guided by implicit knowledge of projective laws (1983), and Ullman's claim that vision transforms polar (i.e., projective) representations into parallel (i.e., non-projective) ones (1979). Even though these theories differ in the computations and depth information used, they all share the view that vision is sensitive to some form of projective information.

**Empirical tests.**

Gibson (1950a; 1950b) was among the first to test the human ability to use texture to judge slant. Subjects matched the slant of textured surfaces, represented in photographs taken from varying viewpoints, by a corresponding inclination of the palm of their hands. Not surprisingly, the patterns of optical texture in the photographs produced responses consistent with an interpretation of a physical surface oriented in depth. However, there was a systematic tendency to underestimate the actual slant -- an effect that was greater for texture patterns that were irregular. Subsequent studies showed that although subjects were not so good at judging absolute slant, they were quite good at judging differences in slant (Attneave, 1972; Beck & Gibson, 1955; Braunstein & Payne, 1969; Clark, Smith, & Rabe, 1955; Epstein, Bontrager, & Park, 1962; Flock, Graves, Tenney, & Stephenson, 1967; Freeman, 1966; Gillman, 1970; Olson, 1974; Perrone, 1980; Phillips, 1970; and Rosinski & Levine, 1976).

Research also showed that relative depth judgments increase in magnitude with increases in simulated depth in texture gradient displays (Stevens, 1981; Todd & Akerstrom, 1987; Vickers, 1971). In Vickers' study, subjects' thresholds of perceived depth were derived from a method of limits procedure where subjects uncovered the
patterns, line by line, until they "received an impression of depth". Immediately afterwards, the pattern was completely exposed, and subjects were asked to bring the slide down from the top until the "impression of depth disappeared". Another experiment in this study, made use of a method of matching and adjustment, whereby subjects indicated the degree of perceived slant by adjusting a corresponding 3-D plane. Similar to Gibson's earlier palm board procedure (1950a), Vickers' study showed that slant and depth impressions become stronger as the number and degree of depth cues increase.

Another experimental technique compared subjects' responses to objects that contain perspective information to those without. The amount of available projective information can be changed simply by manipulating the distance between an object and an image. When the projection onto the image plane is at infinity and therefore parallel, the resulting representation is known as parallel, orthographic (Cutting, 1986), or object-centered (Marr, 1982). When the projection onto the image plane is less than infinity and converges, the resulting form is polar, projective (Cutting, 1986), or viewer-centered (Marr, 1982). The exclusive use of orthographic projection in studies of object perception has been criticized on the grounds that most objects under our scrutiny are relatively near to us (Cutting & Millard, 1984). While perspective projection captures these variations in nearness, orthographic projection simulates infinite viewing distance, a situation approximated only when looking through telescopes at small, distant objects. Thus, in this respect orthographic projection is "unnatural".

Studies that compare subjects' judgments of drawings that either are viewer-centered or object-centered, have shown that we do in fact use projective information in making judgments about the distance and slant of particular objects represented in images (Bengston, Stergios, Ward, & Jester, 1980; Cutting, 1987; Doesschette, 1964;
Gibson, 1979; Hagen, 1980; Johannson & Borjesson, 1980; McGreevy & Ellis, 1986; Nicholls & Kennedy; 1991; Rosinski, Mulholland, Degelman, Farber, 1980). For example, Bengston et al (1980) showed that changing the vantage point affects the perception of distance in accordance with projective geometry. Similarly, Rosinski et al, (1980) tested the effects of vantage point changes on perceived slant. These studies show that the accuracy of distance and slant judgments declines when subjects are positioned at a view other than the depicted vantage point, thus showing our sensitivity to some projective information.

However, there are other studies showing human robustness to deviations from the intended vantage point that counter this view. If we are sensitive to projective information we should be able to detect deviations from it; but subjects in various experiments were unable to detect these deviations (Farber & Rosinski, 1978; Goldstein, 1987; Rosinski et al, 1980). The apparent contradiction of having both sensitivity to projective aspects of an image and robustness to their distortions can be reconciled by any one of three alternative explanations. First, 2-D information from either a surrounding frame or binocular disparity can inform us about the depicted vantage point (Goldstein, 1987). Second, many perspective distortions are indiscriminable from one another (Cutting, 1987). Third, there is evidence that we can compensate for changes in viewing perspective and therefore we may respond independently of viewpoint even though we initially registered viewer-centered information (Hagen, 1980; Perkins, 1973; Pirenne, 1970). Given these accounts of the robustness of perspective, the previously noted tests show that the final human percept uses perspective information in a texture gradient to judge slant or distance.
Chapter 2:

Separating the process from the product

In the review of the literature on 3-D perception in Chapter 1, there was no mention of perceptual processes; responses were simply related to subjective impressions of depth or slant. However, researchers tend to describe this work as though they tested some cognitive or perceptual process (Braunstein, 1976; Stevens, 1981). The danger in extrapolating the results of subjective reports to underlying processes is clearly demonstrated in cases where prior knowledge has been found to be independent of automatic perceptual judgments (Beck, 1966; 1982; Garner, Hake, & Eriksen, 1956; Gillam, 1970; Hochberg, 1956). In Gillam's study, for example, subjects who reported seeing slant performed no differently on a slant matching task from subjects who reported not seeing slant. Similarly, subjects in Beck's study (1966, 1982) showed an incongruity between their judgments of element dissimilarity and texture segmentation. Although a tilted T is judged to be more similar to an upright T than is an L, when repeated to form textures, subjects automatically group Ls with upright Ts, rather than with tilted Ts. While line slope is the more important property for textural segmentation, line arrangement is the more salient property for similarity judgments. Both Beck's and Gillam's experiments warn of generalizing findings from verbal reports to performance under stress of time or accuracy. Different factors and perceptual processes may influence rapidly-made versus contemplative decisions.

The few experiments that are an exception to the large number of studies employing depth or slant ratings include speeded performance tests (Bennett & Warren; 1993; Leibowitz & Bourne, 1956; Pringle & Uhlarik, 1982; Smets & Stappers; 1990; Uhlarik, Pringle, Jordan, & Misceo; 1980), and tests that manipulate attention (Epstein & Babler; 1989, 1990; Epstein, Babler, & Bownds, 1992; Epstein & Broota, 1986; Epstein & Lovitts; 1985). These are the only documented psychophysical studies, to
date, studying the visual process(es) involved in the perception of depth from perspective information. In Epstein & Broota (1985), subjects compared the size of objective and projective representations of objects to previously encountered figures. Shapes rotated in depth were presented in conditions that differed in the amount of attention directed to the matching task. In trials where subjects' attention was fully devoted to the task, they chose the objective alternative as a match for the standard. However, in trials where attention was diverted to an odd-even discrimination task, subjects chose the projective alternative. Epstein and Broota use these findings as evidence for two operations being used in size-distance processing (See also Epstein & Lovitts; 1985). First, there are automatic processes that register projective size and shape, and second, there is an attentive process that integrates those outputs into an object-centered description. Additional tests of this theoretical account (Epstein & Babler; 1989; 1990; Epstein, Babler, & Bownds, 1992) are presented in Chapters 2 and 5.

Rival accounts show that the allocation of attention influences the effectiveness of depth cues (Coren & Porac, 1983; Gogel, 1967; Gogel, Loomis, Newman, & Sharkey, 1985; Gogel & Tietz, 1976; Peterson, 1986; Shulman, 1991a; 1991b; Tsal, 1984). However, these studies use illusion magnitude assessments that do not isolate depth information from alternative mechanisms known to mediate illusions (i.e., test and inducing element confusion, element averaging, size-contrast). Therefore, the extent to which these findings generalize specifically to the perception of depth is unknown.

Further ambiguity as to the object specified by the image is introduced by the impoverished nature of the stimuli in these and all of Epstein's studies (i.e., availability of only edge or isolated line information). By contrast, a textured surface tends to provide information about surface material, slant, curvature, as well as edges. The speeded performance tasks noted above come the closest to assessing underlying visual
processes involved in the interpretation of textured surfaces. Pringle and Uhlarik (1982) and Bennett and Warren (1993) used a simultaneous matching RT task to determine if shape (and size) recognition depended on environmental or retinal size. They used a display containing a textured hallway in the background to manipulate these two representations of size. Display exposures which were terminated by subjects' responses ranged between 1100 and 1500 msec. Both environmental and retinal information contributed to matching performance. The primary interest of these studies was to assess underlying representations as indicated by their environmental/retinal manipulation and to determine if size-scaling utilized a similar process to the one used in mental rotation (i.e., Larsen & Bundeson, 1978; Rock & Linnett, 1993). Use of relatively brief exposure manipulations and indirect assessments of perceived distance via shape matching ensured that they were measuring performance based on depth perception rather than subjective impressions. However, since exposure duration ranged between 1100 and 1500 msec, these effects are still likely to be mediated by attention. (Pringle and Uhlarik found similar effects for RTs which ranged between 900 and 4000 msec.) Nevertheless, these studies do show that subjects interpret apparent depth from texture gradient information within 1500 msec on a task that directly tests performance speed rather than phenomenal depth perception.

Stevens (personal communication) and Smets and Stappers (1990) attempted to further isolate early apparent depth processing by limiting exposure duration to texture gradient displays. Stevens has studied the task of quantitatively matching depth between random dot stereograms and monocular renderings of surfaces depicted by a grid of contours. In one set of conditions, the random dot stereogram depicted a gaussian form in depth (viewed such that the bump protrudes in stereo depth towards the viewer). An exposure of 100 msec, was sufficient to allow subjects to match by adjusting the amplitude of the subsequently viewed stereoscopic gaussian with the
apparent depth of the monocular gaussian. In the reverse situation, as little as 70 msec of presentation of the monocular depiction gave sufficient 3-D information for subjects to accurately match the two gaussians.

Smets and Stappers (1990) also tested (100 msec) brief exposure to a textured display, but here the subjects' task was to detect a texture element oriented opposite to the underlying texture gradient. Detection depended on disrupting the texture gradient pattern and was interpreted as evidence for early processing of the high-order relation information (i.e., first or second derivatives) present in a texture gradient as opposed to the reliance of early perception on simple local features (2-D orientation). However, Smets and Stappers did not adequately separate these two forms of information. The use of a gradient of items as both the search items and surrounding context resulted in an inadequate control for local differences in 2-D orientation. Consequently, their effect may have been mediated either by simple orientation discriminations between the target and background lines or by 3-D slant information.

The present study avoids this methodological confound by using search items that are manipulated separately from the slanted context and that are constant in size across the gradient. A second improvement includes varying the number of elements in the visual search display. By relating response times to the size of display we can make stronger inferences about the nature of visual processing, as described in Chapters 3 and 4.

**Do subjective measures of slant-from-texture predict speeded performance?**

The first objective of the present study is to determine if slant-from-texture can influence visual search performance. The subjective rating literature assures us that geometric projective rules are used at least in a number of subjective rating tasks (See Chapter 1). Since the majority of these experiments allow unlimited time for subjects to respond, the consequence may be that these tests have measured only the output of
perceptual processing. Such testing may not reflect our more immediate and spontaneous visual processes. Only the few noted experiments restrict stimulus exposure typically to about 100 msec to test depth perception. However, many are subject to confounding influences (Epstein, Babler, & Bownds, 1992; Epstein & Broota, 1986; Epstein & Lovitts, 1985), or remain vague about underlying processes (Epstein & Babler, 1989; 1990; Smets & Stappers, 1990). Therefore, it remains unclear how soon projective rules are spontaneously used in visual processing and how we should characterize such processes.

The visual search task

In the visual search experiments presented here, one critical experimental manipulation was search for an item on an apparently slanted surface as compared with search for an item on a control surface containing no slant. The slanted surface was depicted at an angle of 77° from the viewer's frontal plane, and the control surface was parallel to the viewer's frontal plane as is shown in Figure 4. Subjects were required to search for an item that was unique in size (i.e., short target among long distractors or long target among short distractors), and located on one of the two surfaces (77° or 0°). The prediction that visual processing of these items is influenced by apparent depth information from a texture gradient will be supported if search performance on the control surface differs from search on a slanted surface in a manner consistent with the depicted depth information.

The pattern of the effect is predicted most simply by the projective size relation -- an item in the distance appears retinally smaller than the same item up close. Therefore, visual search is expected to be easier when the long target is far, and harder when it is near. The inverse performance is expected for the short target -- search will be easy when the short target is near and difficult when it is far. Such a contextual influence from the background texture gradient will counter the prevailing view that rapid visual
processing depends only on simple image features such as length, orientation, and color (See Chapters 3 and 5).

Take note that the texture gradient studied here differs from the depth information used in most of the previously noted experiments (i.e., Epstein & Babler; 1989, 1990; Epstein, Babler, & Bownds, 1992; Epstein & Broota, 1986; Epstein & Lovitts; 1985), as well as conventional visual search studies (Treisman & Gormican, 1988). Epstein et al. manipulated the presence of depth information in an object; here depth information is manipulated in a surface against which the objects are arranged. Nevertheless, Epstein, Babler and Bownds (1992) point out the importance of having search items presented in an arrangement that fosters an impression of a surface. In their study, subjects were much faster in detecting a rotated target among frontal-parallel distractors as compared with detecting a frontal-parallel target among rotated distractors. This performance asymmetry reflects the greater signal-to-noise ratio elicited by an interpretation of a continuous surface of non-rotated distractors being disrupted by a rotated target. The relevance of this finding to the present study is that it shows the importance of background coherence. Perhaps search would be just as rapid in the case of detecting a frontal-parallel target among rotated distractors if rotated distractor items, as a group, shared a common vanishing point.

Another reason to present perspective information in the form of a surface rather than as a small object is to show realistic and discernible convergence. Realistic convergence occurs when the distance from the center of projection to the surface (See Figure 2) allows between 30° to 60° of the scene to be shown (Cutting, 1986; 1987; Sedgwick, 1986; 1987). Often a surface (or a region of the surface) will fit this range, whereas a small item like those used in visual search experiments, typically takes up only a few degrees. If the width of the scene is too narrow (< 30°), the depth of perspective becomes flattened, and if it is too wide (> 60°), the perspective appears
exaggerated (Walters & Bromham; 1970). Placing the image plane between the center of projection (i.e., reducing the projected image to < 30°) is sufficient as long as the relative convergence remains the same. Pilot testing indicated that at least 5° of visual angle is required for convergence of edges on an object to be detected. This constraint was satisfied in this study by using a representation of a 45° textured surface reduced 33% to match the size of the computer screen (14° by 19° of visual angle).

To summarize, the rationale for placing perspective information in a surface rather than in the search items includes: (1) Using the texture gradient as the background helps avoid the potential confounds that occurred in Smets and Stappers' study from using a gradient of distractors; (2) A textured surface is likely to foster more background coherence; (3) The 30° to 60° “realistic convergence” range constraint is satisfied; (4) Subjects are able to resolve the converging line information.

As noted earlier, evidence for depth processing in visual search will be based on influences from size and location. Obtaining any of the predicted effects (i.e., more difficult search when the short target is far or the long target is near) will show that visual search is sensitive to 3-D information and the mediating mechanism accordingly scales for projective size.

**Experiment 1:**

**Slant-from-texture influences on visual search**

This first experiment used the size discrimination task described above where search items were presented on a textured pattern spanning the entire display screen. The two-by-two design consisted of a slant versus control condition, and a long target - short distractor condition versus a short target - long distractor condition. Search items were depictions of realistic cylinders that appeared to rest on the underlying slanted surface.

Realistic object-like items were selected because simple line items resembling those used by Treisman and Gormican (1988) had unreliable effects across locations in
pilot testing. The textured surface may have only elicited undifferentiated noise in slant and control conditions. The simple line search items may have also allowed subjects to ignore the background information since it had the appearance of being independent of the search task. In order to increase the likelihood that search strategies would be influenced by the surrounding context in a manner consistent with the depicted depth information, more object-like items were presented on the display surface.

Method

Subjects. Eleven members of the University community participated in two 60-minute sessions to complete eight conditions of the visual search experiment. Each condition consisted of 3 sets of 60 trials. Nine subjects had no experience with visual search tasks. The remaining two were experienced with the task but did not have specific experience with the search items or the background texture used here. Subject age ranged from 17 to 33 years; eight were female and three male. All subjects had normal or corrected to normal vision.

Stimuli and Procedure. Display presentation and data collection were controlled by a Macintosh computer using the MacLab program (Costin, 1988). The subjects' task was to search for a target item among 2, 6, or 10 items. All of the search items were cylinder-like in appearance and were randomly distributed on a 19°X 14° display screen. Targets differed from distractors only in size and were present on a random one-half of the trials.

Targets had an overall image extent of 1.35 X 1.0 cm, subtending 1.5 x 1.1°. The cylinder's width and wall height were 1.0 cm in the image. Distractors in the first two conditions had an overall extent of 1.7 X 1.2 cm, subtending 1.8 x 1.2°. The distractor cylinders' width and wall height were 1.1 cm. Distractors in the second two conditions had an overall extent of 1.2 X .9 cm, subtending 1.4 x 1.03°, and a cylinder wall height and width of .9 cm. In each of the blocks, targets were presented four times
in each of the three locations: upper, middle, and lower visual fields. Search items were initially selected from a set of cylinders that were placed along the line of sight of a 77° slanted surface that contained a systematic texture similar to the one shown in Figure 4. Distractors in the first slant and control conditions were larger than the target by two texture element steps toward the viewer, while distractors in the second two conditions were smaller by an equivalent two steps away from the viewer (and target).

Textured surfaces used in the visual search displays were generated by the Dynaperspective Design and Modeling program (Tatsumi & Okamura; 1988). The display screen located 50 cm from the viewer, contained regularly textured surfaces that appeared to be 80 cm in front of the viewer (apparently 30 cm behind the computer screen and spanning 45° of visual angle). The control grid was a regular two-dimensional grid containing .9 x .9 cm grid elements on the computer display and had no slant relative to the viewer's frontal plane. The texture gradient used in the Experimental condition was a depiction of the control grid slanted at a 77° angle as is shown in Figure 4. The gradient was oriented vertically with the densely spaced texture elements depicting far located at the top of the screen, and the more diffusely spaced elements depicting near at the bottom. The gradient consisted of horizontal components which ranged from 1.1 to 2.5 cm from the top to the bottom of the screen, with a mean width of 1.7 cm. The gradient also consisted of vertical components which ranged 0.25 to 1.5 cm from the screen's top to its bottom, with a mean height of .7 cm. All texture outlines were black (12.2 cd/m²) and all backgrounds were white (158.9 cd/m²).

Figure 4 shows a sample display where the task is to search for a short cylinder against a background of longer cylinders.

Each trial began with a fixation symbol lit for 500 ms, followed by the display which remained visible until the subject responded. Target presence and absence were
Figure 4. The texture gradient used in visual search conditions was a depiction of a vertical and horizontal grid slanted at a 77° angle with densely spaced texture elements depicting far, and more diffusely spaced elements depicting near. Cylinder items resting on the textured surface were used as the target and distractor items in the search task.
reported by pressing one of two response keys. Accuracy feedback was presented in the form of a beep on incorrect trials. Subjects were instructed to maintain fixation throughout the trial sequence, to respond as rapidly as possible, and to keep errors at a minimum.

Results

Mean correct response time (RT) and percentage of errors for target present trials are shown in Figure 5; standard errors of the means are shown in Table 1. Target absent trials were not examined in detail since they could not be differentiated by location. The sensitivity of present trials is described below in a comparison of performance across the locations of the slant and control conditions. Reported statistics are based on the overall ANOVAs which include the two level factor of target presence and the three level factor of target location.

Overall, search was more difficult when the short target was in the top-far location relative to the middle and bottom-near locations of the slanted display (RT and errors: Fisher's LSD tests, $p < .01$). Mean RT was 386 msec slower and errors were 27% greater when the short target was far (mean RT = 1230 msec, mean error = 36%) relative to when it was near (mean RT = 844 msec, mean error = 9%).

As expected, the inverse trend occurred when subjects searched for a long target against short distractors. Search was more difficult when the long target was in the bottom-near location relative to the middle and top-far locations (RT and errors: Fisher's LSD tests, $p < .001$). Mean RT was 387 msec slower and there were 33% more errors when the long target was near (mean RT = 1220 msec, mean error = 43%), than when it was far (mean RT = 833 msec, mean error = 9%).

Surprisingly, slightly worse performance occurred when the short target was at the top of the control condition (RT: Fisher's LSD tests, $p < .05$; errors: Fisher's LSD tests, $p < .05$).
Figure 5. Mean correct reaction time and percentage of errors for target present trials of Experiment 1. Two search items (short and long) were examined against two backgrounds (slant and control) and three display sizes (two, six and ten). Critical comparisons are between locations in the slant background (far, middle and near), and the control condition (top, middle and bottom). SEMs are reported in Table 1.
Table 1

<table>
<thead>
<tr>
<th>Display Size</th>
<th>Apparent Slant Location</th>
<th>RT (msec)</th>
<th>Standard Error of the Mean (SEM) in Experiment 1</th>
<th>Errors (%)</th>
<th>SEM</th>
<th>SEM</th>
<th>SEM</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slant</td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Far</td>
<td>86</td>
<td>70</td>
<td>7.3</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>56</td>
<td>54</td>
<td>5.4</td>
<td>5.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>54</td>
<td>48</td>
<td>4.9</td>
<td>9.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Far</td>
<td>101</td>
<td>121</td>
<td>6.1</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>71</td>
<td>46</td>
<td>4.1</td>
<td>4.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>43</td>
<td>33</td>
<td>5.5</td>
<td>7.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Far</td>
<td>119</td>
<td>114</td>
<td>7.6</td>
<td>4.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>97</td>
<td>91</td>
<td>0.0</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>57</td>
<td>50</td>
<td>5.7</td>
<td>5.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slant</td>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SEM</td>
<td>SEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long Target</td>
<td>Far</td>
<td>33</td>
<td>33</td>
<td>6.5</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>27</td>
<td>41</td>
<td>4.3</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>77</td>
<td>55</td>
<td>2.5</td>
<td>5.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Far</td>
<td>64</td>
<td>25</td>
<td>5.8</td>
<td>3.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>56</td>
<td>35</td>
<td>3.3</td>
<td>2.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>101</td>
<td>67</td>
<td>2.6</td>
<td>5.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Far</td>
<td>66</td>
<td>35</td>
<td>8.0</td>
<td>6.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>80</td>
<td>49</td>
<td>0.3</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>126</td>
<td>83</td>
<td>3.3</td>
<td>5.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
tests \( p < .001 \) and the long target was at the bottom of the control condition (RT: Fisher's LSD tests, \( p < .001 \); errors: Fisher's LSD tests \( p < .001 \)). RT was 145 msec slower with 15\% more errors when the short target was at the top of the display (mean RT = 1011 msec; mean error = 18\%) relative to the bottom (mean RT = 866 msec; mean error = 4\%), and RT was 122 msec slower with 14 \% fewer errors when the long target was on the top (mean RT = 840 msec, mean \% error = 8\%) relative to the bottom of the display (mean RT = 961 msec; mean error = 22\%).

Although these location effects in the control condition were half the magnitude of those in the slant condition they were significant on the noted pairwise comparisons. This inverse relationship between short and long target performance in the slant and control conditions was also confirmed by the significant interaction between target location and target size for RT, \( F(2,20) = 40.0, p < .001 \); and errors, \( F(2,20) = 21.1, p < .001 \).

However, the stronger interaction between target size and target location in the slant condition relative to the control condition was supported by a significant 3-way interaction between background, target location, and target size (RT: \( F(2,20) = 14.1, p < .01 \); errors: \( F(2,20) = 19.1, p < .001 \)), and a significant 4-way interaction between background, target location, target size and target presence (RT: \( F(2,20) = 17.6, p < .001 \); errors: \( F(2,20) = 7.2, p < .01 \)). The latter interaction simply indicates that apparent location effects were found only in target present trials.

Sensitivity to differences in target location for target present trials is further reflected in target presence interaction with target location (RT: \( F(2,20) = 8.6, p < .01 \); errors: \( F(2,20) = 11.4, p < .001 \)), target size (RT: \( F(2,20) = 8.2, p < .05 \)), and size and location (3-way ANOVA -- RT: \( F(2,20) = 37.0, p < .001 \); errors: \( F(2,20) = 23.6, p < .001 \)). Target present trials also account for the main effect for target location (RT: \( F(2,20) = 30.4, p < .001 \); and errors: \( F(2,20) = 22.3, p < .001 \)). Additional
interactions comparing target presence and target absence are described in Appendix A.

**Display size effects.**

There was a substantial increase in the range of RTs across display sizes in target present trials, when the short target was far and the long target was near (RT: display size, target size and target presence -- $F(2,20) = 6.4, p < .01$; and display size, target size and target location, $F(4,40) = 6.9, p < .001$). Two additional interactions show that this trend was stronger in slant conditions (display size, background, X target presence, $F(2,20) = 5.9, p < .05$; and display size, background, target presence, target size X target location, $F(4,40) = 7.7, p < .001$).

**Central location effects.**

An additional trend that is readily apparent in Figure 5 is the better performance that occurred when the target was located in the central region (All long target trials RTs: Fisher's LSD tests, $p < .05$). This facilitation effect is likely the result of surrounding distractors acting to ease the size-discrimination task. This interpretation is supported by item density influences on general visual search performance (Banks, Larson, & Prinzmetal, 1979; Boynton, Hayhoe, & MacLeod, 1977), and Gathercole and Broadbent's (1987) finding that the effects of distractors depend on their distance from the target. Several additional studies show that search accuracy can improve with more distractors (Green, 1992; Sagi, 1990, Sagi & Julesz, 1985, 1987). Sagi and Julesz interpret the latter effect as support for preattentive search making local comparisons in a feature gradient. More simply we can regard the effect as a consequence of facilitation from proximity -- as the items get closer together, comparisons become easier.

**Location and testing order.**

Surprisingly, location influences similar to those found in the slant condition occurred in the control conditions. Performance was easier when the short target was at the bottom of the display and the long target was at the top. One potential cause for
these top-to-bottom effects may involve depth information from the slant conditions carrying over into the control conditions. Even though slant and control trials were presented in separate blocks, and the order of presentation was counterbalanced across subjects, it is possible that a strong initial order effect could influence the data. Depth information carry over could only occur in one direction -- from slant to control, since depth information is present only in slant conditions. To test this hypothesis subjects were separated across first condition received: long or short target, and slant or control. No relevant interactions, nor main effects involving order was significant. The only significant interactions included presentation order X location for RT, $F(6,14) = 3.4, p < .05$, reflecting better performance in the top locations for those subjects who received the control conditions first. Order of presentation therefore cannot account for the top-bottom trends that occurred in the control condition.

**Discussion**

Performance in the slant condition was consistent with the texture gradient eliciting an apparent depth effect -- search was more difficult when the short target was far and the long target was near. However, smaller but similar effects occurred in the control condition -- search was more difficult when the short target was at the top of the display and the long target was at the bottom. Faster search for both conditions in the central region is easily explained by surrounding distractors providing comparison stimuli to ease discrimination. What accounts for the unexpected performance in the control condition? Since testing order was ruled out, we are left with some top-to-bottom influence, perhaps differential sensitivity to information in the top versus the bottom of our visual field, depending upon the size of target.

Another possible account is that these locations are interpreted by subjects as providing height-in-the-plane information, where items in the top region appear to be farther away than items in the bottom region. This interpretation is consistent with Sedgwick's
(1986) demonstration that a picture reveals an object's size relative to the height of the plane even without explicit horizon information being present. Moreover, Bruno and Cutting (1988) have shown such height-in-the-plane effects on performance in a magnitude estimation and depth rating task. In a further attempt to isolate apparent distance effects arising solely from the texture gradient, the visual search task in Experiment 2 separated top-to-bottom information from apparent-distance information in the texture gradient.

Experiment 2:

Separating top-to-bottom from apparent-distance effects

In Experiment 1, performance was the poorest when the short target was located at the top of the display, and the long target was located in the bottom. This was true for both slant and no slant conditions. This size-location influence on visual search in the control condition may have been caused by either of two factors: (1) different discrimination capabilities across the visual field, or (2) inferred depth from height-in-the-plane. Since performance strongly resembled predicted projective trends in the control condition, a likely explanation is that subjects infer height-in-the-plane depth information from top-to-bottom cues.

One way to separate top-to-bottom from apparent distance influences is to change the orientation of the background surface texture. By doing so, display directions (i.e., top, bottom, left, right) and apparent depth (i.e., near, far) are no longer confounded. In this experiment, visual search is performed in displays where the textured surface appears to slant sideways. Now instead of the texture gradient appearing as a textured floor (as in Experiment 1), the texture gradient appears as a textured wall that is slanted to the left or to the right. Because top-to-bottom information is no longer confounded with near-to-far information, we can isolate apparent distance effects arising solely from the texture gradient.
Method

Subjects. Eleven University students participated in two one-hour sessions to complete 3 sets of 60 trials for each of eight conditions. Eight subjects were inexperienced in visual search tasks. Subject age ranged from 17 to 27 years; six were female and five were male. All subjects had normal or corrected to normal vision.

Stimuli and Procedure. All stimuli and procedures were the same as those used in Experiment 1. The only change involved rotating the display 90° to the left for half of the trials, and 90° to the right for the other half of the trials. Left-right orientation was counterbalanced across subjects to prevent lateral search biases that may arise in visual search performance.

Displays were again subdivided into three regions to allow for an analysis of the effects of target location on visual search performance. In the slant condition, the three regions were designated as far, middle, and near. In the control condition, these regions corresponded to left, middle, and right when the display was rotated to the left, and right, middle, and left when the display was rotated to the right.

Results

Mean correct RT and percentage of errors for target present trials are shown in Figure 6; standard errors of the means are shown in Table 2. Since only target present trials were sensitive to differences in target location, the results focus on these trials with a direct comparison of performance across the extreme locations of the slant and control conditions. Performance here, as in Experiment 1, was affected by target location and size. However, in this experiment, size-location effects were different for the slant and control conditions.

--- Insert Figure 6 & Table 2 ---

In the slant condition, search was more difficult when the short target was in the far location relative to the middle (RT and errors: Fisher's LSD tests, p < .01) and the
near locations (RT and errors: Fisher's LSD tests, $p < .01$). RTs were 262 msec slower and errors 20% greater when the short target was far (mean RT = 1165 msec; mean error = 26%) than when it was near (mean RT = 903 msec; mean error = 6%). This contrasts with the control condition where RT and errors were not significantly different on either side (far control: mean RT = 897 msec, and mean error = 8%; near control: mean RT = 895 msec, and mean error = 7%).

The expected inverse relation between performance and location emerged when subjects searched for the long target against short distractors on the slanted background. Search was more difficult when the long target was in the near location relative to the middle (RT and errors: Fisher's LSD tests, $p < .01$) and far locations (RT and errors: Fisher's LSD tests, $p < .01$). Here RT was 379 msec slower and errors were 30% greater when the long target was near (mean RT = 1253 msec; mean error = 35%) than when it was far (mean RT = 874 msec; mean error = 6%). This contrasts with performance in the control condition where RT and errors were not significantly different on either side (near control: mean RT = 1108 msec, and mean error = 16%; far control: mean RT = 1014 msec, mean error = 11%).

The inverse trend between the short and long target in the slant condition was confirmed by a significant 2-way interaction between target location and target size for RT, $F(2,36) = 30.6$, $p < .001$ and errors, $F(2,34) = 46.3$, $p < .001$, as well as the significant interaction between background and target size (RT: $F(1,18) = 11.0$, $p < .01$). Separate main effects occurred for target location (RT: $F(2,36) = 82.9$, $p < .001$, and errors: $F(2,34) = 35.9$, $p < .001$), target size (errors: $F(1,17) = 7.4$, $p < .05$), background (RT: $F(1,18) = 5.9$, $p < .05$; and errors: $F(1,17) = 13.5$, $p < .001$); and target presence (RT: $F(1,18) = 49.5$, $p < .001$; and errors, $F(1,17) = 26.7$, $p < .001$). Additional interactions comparing target presence and target absence trials are noted in Appendix A along with an analysis of location using target absent trials as a control.
Figure 6. Mean correct reaction time and percentage of errors for target present trials of Experiment 2. Two search items (short and long) were examined against two backgrounds (slant and control), and three display sizes (two, six and ten). Critical comparisons are between locations in the slant background (far, middle and near), and in the control condition (top, middle and bottom). SEMs are reported in Table 2.
<table>
<thead>
<tr>
<th>Display Size</th>
<th>Apparent Location</th>
<th>RT (msec)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slant SEM</td>
<td>Control SEM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short Target</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Far</td>
<td>63 38</td>
<td>5.9 2.1</td>
<td></td>
</tr>
<tr>
<td>2 Mid</td>
<td>48 20</td>
<td>2.2 2.2</td>
<td></td>
</tr>
<tr>
<td>2 Near</td>
<td>31 41</td>
<td>1.2 1.7</td>
<td></td>
</tr>
<tr>
<td>6 Far</td>
<td>70 35</td>
<td>5.2 1.2</td>
<td></td>
</tr>
<tr>
<td>6 Mid</td>
<td>56 26</td>
<td>2.7 3.2</td>
<td></td>
</tr>
<tr>
<td>6 Near</td>
<td>31 26</td>
<td>2.2 2.2</td>
<td></td>
</tr>
<tr>
<td>10 Far</td>
<td>71 51</td>
<td>4.9 4.0</td>
<td></td>
</tr>
<tr>
<td>10 Mid</td>
<td>57 59</td>
<td>0.3 0.4</td>
<td></td>
</tr>
<tr>
<td>10 Near</td>
<td>45 44</td>
<td>3.0 3.8</td>
<td></td>
</tr>
<tr>
<td>Long Target</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Far</td>
<td>33 24</td>
<td>1.8 1.7</td>
<td></td>
</tr>
<tr>
<td>2 Mid</td>
<td>34 28</td>
<td>2.8 3.2</td>
<td></td>
</tr>
<tr>
<td>2 Near</td>
<td>59 40</td>
<td>5.8 2.9</td>
<td></td>
</tr>
<tr>
<td>6 Far</td>
<td>29 41</td>
<td>2.0 1.7</td>
<td></td>
</tr>
<tr>
<td>6 Mid</td>
<td>35 28</td>
<td>2.8 2.2</td>
<td></td>
</tr>
<tr>
<td>6 Near</td>
<td>63 53</td>
<td>4.5 0.0</td>
<td></td>
</tr>
<tr>
<td>10 Far</td>
<td>42 41</td>
<td>3.6 1.8</td>
<td></td>
</tr>
<tr>
<td>10 Mid</td>
<td>44 35</td>
<td>0.0 3.9</td>
<td></td>
</tr>
<tr>
<td>10 Near</td>
<td>61 56</td>
<td>5.6 4.9</td>
<td></td>
</tr>
</tbody>
</table>
Discrepancies in degrees of freedom for RT and Error data are due to software errors in identifying subjects' data.

**Location effects in the control condition.**

Performance was easiest when the target was in the middle region. This trend was supported by a significant main location effect on simple effects test of the control condition for the long (RT: $F(2,42) = 20.1, p < .001$; and errors: $F(2, 36) = 9.8, p < .001$), and short target trials (RT: $F(2,42) = 17.4, p < .001$). Fisher LSD tests confirmed that performance improved when the long target was located in the central region (All RTs and errors, $p < .05$). For the short target, performance did not significantly differ across locations in the control condition (All Fisher's LSD tests, $p > .05$). Only when separated across display size did pairwise comparisons reach significance (i.e., ten item condition; see display size effects).

Unlike the top-to-bottom effects in Experiment 1, performance did not differ between the left and right locations of the control condition (All Fisher's LSD tests, $p > .05$). An analysis of left versus right orientation as a between factor revealed no difference between these orientations for RT ($F(1,17) = .26$), and errors ($F(1,16) = .45$), indicating that rotating the display was successful in controlling influences from top-to-bottom search biases.

**Display size effects.**

RTs showed the expected increase in display size across all conditions, $F(2,36) = 114.4, p < .001$. The widest range of RTs across display sizes occurred in the far location of the short target trials, and surprisingly in the extreme locations of the long target control condition. These trends were supported by the three-way RT interaction between display size, background and target size, ($F(2,36) = 6.6, p < .01$), the five way interaction between target presence, display size, background, target size, and location for RT, ($F(4,72) = 5.0, p < .001$), and the additional interactions involving
display size noted in Appendix A. RT slope analysis is described further in Figure 7 and Chapter 3.

Testing order.

It is clear from the performance trends in the control conditions that testing order was not responsible for location effects. Nevertheless, the absence of an order effect was confirmed by a between-subjects ANOVA that separated subjects into four groups differing on the first condition received: long or short target, and the presence or absence of slant. Presentation order was not significant for both RT, $F(3,6) = .13$; and errors, $F(3,6) = .16$. The only significant interaction involved presentation order x slant for RT ($F(3,6) = 5.3, p < .05$) showing a trend in the control condition inconsistent with those in the slant condition.

Discussion

Experiment 2 established that the speeded performance task is reliable in assessing our ability to access depth information from a texture gradient. Evidence for depth processing was revealed in the slant condition -- search performance was poorest when the short target was far, and the long target was near (mean RT = 1215; mean errors = 28%), and search performance was best when the long target was far and the short target was near (mean RT = 889 msec; mean errors = 5%). These effects disappeared when information about distance was removed in the control condition (mean RT = 979 msec; mean errors = 8%). The only remaining effects included the expected central region improvement from the presence of surrounding distractors.

Evidence for depth processing in the speeded visual search task suggests that neither subjective judgments of depth, nor acts of volition drive the effect. Presumably contemplative judgments are at a minimum in a task that avails the subject only about one second of viewing (See RTs in figure 6). Nonetheless, apparent depth and assumptions of projective size are evident in such rapid perceptual judgments. The
remaining experiments use this visual search method to further clarify how early vision is sensitive to apparent depth in the texture gradient, and which texture dimensions are the most informative in rapid perception.
Chapter 3: How early are the slant-from-texture influences?

A two stage model of vision

Contemporary psychophysical accounts of visual processing argue that vision consists of two systems that may be defined in terms of (1) overall response time, or (2) visual search strategies. These two operational definitions allow us to make a further theoretical distinction between the processing complexity of early and late vision.

The first and simplest definition of the visual subsystems is based on absolute exposure time, where early vision occurs within 100 msec of the display onset and late vision requires at least 100 msec (Treisman, Cavanagh, Fischer, Ramachandran, and von der Heydt, 1990). This rough delineation between early and late vision is typically determined by restricted exposure experiments where subjects accurately detect an item or group of items within the 100 msec of allotted time. The two systems, although sometimes described as dichotomous, are best regarded as continuous, given that tests have shown gradually changing accuracy with exposure (Callaghan, 1989; Duncan & Humphreys, 1989; Northdurft, 1985; Treisman & Souther, 1985).

The second, definition is based on visual search and focuses on the ability to perform simultaneous processing of items. Here, parallel processing is spatial, rather than temporal. According to Treisman's Feature Integration Theory (1986) the first preattentive system detects features in parallel across the visual field, and the second attentive system sequentially processes the relations between these features. In other words, the preattentive system initially encodes the features, and the attentive system combines the features to form object representations (Beck, 1982; Julesz, 1984; Treisman, 1986). Researchers typically operationalize preattentive and attentive systems in terms of search rate as a function of the number of items in the display (i.e., RT slope). Slopes falling below 10 ms/item are usually classified as a parallel search and those above are considered serial. Although recent evidence suggests that the
preattentive and attentive system lie on a continuum of processing strategies (Duncan & Humphreys, 1989; Treisman & Souther, 1985), for the purpose of exposition, reference will be made to two systems with the implication that these are graded categories of visual processing.

An important theoretical distinction between early and late vision is derived from the previous operational definitions. This distinction focuses on the complexity of information that each system can code. Data show that the earlier preattentive system codes simple geometric features, such as color, orientation, length, and curvature, and the later attentive system operates on the conjunction of features (Beck, 1982; Julesz, 1984; Kröse, 1987; Treisman & Sato, 1990). However, recent evidence suggests that early vision also may be capable of registering somewhat more complex information (Treisman & Gormican, 1988; Duncan & Humphreys, 1989; Enns & Rensink, 1990, 1991; Humphreys, Quinlan, & Riddoch, 1989; Ramachandran, 1988). These findings have prompted reevaluation of the function of each stage, and our conception of primitive features. One proposal is to expand early vision's sensitivity to include various combinations of features that correspond to properties of the scene (Enns, 1992). For example, rapid search for conjunctions of features such as binocular disparity and motion (Nakayama & Silverman, 1986), or motion and shape (McLeod, Driver, & Crisp, 1988), may reflect early sensitivity to apparent depth. Similarly rapid search for spatial relations among lines and shaded polygons suggests early sensitivity to 3-D orientation and the direction of lighting (Enns & Rensink, 1990a; 1990b; 1991).

The scene-based properties of the texture gradient can be described in a variety of ways including the rules of projective geometry noted earlier that relate surface slant to the texture variations in the image. An analysis of this scene-based information, its corresponding image information, and their relation to early vision is presented in the final discussion.
Encoding slant-from-texture: Early or late?

As noted in Chapter 3, the few published experiments that test early processing of projective information have restricted stimulus exposure to 150 msec or less (Epstein & Babler, 1990; Smets & Stappers, 1990), or have tested preattention by diverting attention to a secondary task (Epstein & Broota, 1986; Epstein & Babler, 1989). These studies, along with a recent visual search experiment comparing search for a projective versus objective shape (Epstein, Babler, & Bounds, 1992), show that early vision is sensitive to projective information. Certain qualifications are required in interpreting these findings. First, in all of these studies, projective information was confounded with additional cues to depth. In Epstein et al's studies, projective information was confounded with binocular disparity and shading information, and in Smets and Stappers' (1990) study texture gradient information was confounded with simple orientation information. Second, only Smets and Stappers and Stevens actually used texture information. Epstein et al's experiments presented projective information only through outlining contours. Perceptual sensitivity to projective information may differ for surrounding contour and surface-texture information. Given the research findings to date, it remains unclear how early slant-from-texture is registered in visual processing.

One way to test early vision's sensitivity to slant-from-texture is to utilize a conventional test that distinguishes early and late vision. In visual search tasks, early and late vision are distinguished by how RTs are affected by the number of items in the display. The system's sensitivity to information is determined by the further interaction with the kind of search items. Early sensitivity is shown if visual search is eased to a point where a slow-serial search is speeded up to a fast-parallel strategy when either the short target is near or the long target is far.
RT Slope performance

Figure 7 summarizes RT slope performance from Experiment 2 across near and far locations of target present trials. The mean RT slope across locations is 30 ms/item indicating that performance in all conditions requires the use of attention. Figure 7 also shows a trend towards faster search when the short item was near (Fisher LSD’s, \( p < .05 \)). Although, this trend is supported by the 4-way interaction between background, target size, target location, and target presence (\( F(2,36) = 6.7, p < .05 \)), there is no similar difference across locations for the long target.

Additional RT slope effects presented below clarify the above noted trend. RT slopes were slightly smaller in slant relative to the control conditions (\( F(1,18) = 7.7, p < .05 \)). This difference is a consequence of target location interacting with target size in slant conditions (\( F(2,36) = 4.6, p < .05 \)). The only effect from location in the control condition was an improvement when the target was in the middle region. This trend was supported by a significant main location effect (\( F(2,36) = 9.5, p < .001 \)), as well as simple effects tests that showed the bulk of the effect on long target - control trials (\( F(2,57) = 9.9, p < .001 \)). Search for the short target was not significantly different across locations in the control condition (All Fisher’s LSD tests).

Projective size or apparent proximity?

RT slopes do not appear to be sensitive to slant-from-texture in long target trials. Yet, RT slopes for short-target trials do appear sensitive to apparent depth information. Perhaps in addition to a projective size effect arising from the textured gradient, there may be an influence from the apparent proximity of the search items. The target that appears closer, regardless of its size, may simply be easier to detect than a further target. Fast target present RT slopes in the near location (23 msec/item) relative to far (39 msec/item) suggest this may be so.
Figure 7. Mean correct reaction time slopes (ms/item) in Experiment 2. Two search items (short and long) were examined against two backgrounds (slant and control) and two locations (far and near in the slant condition, and left and right in the control condition).
Such a proximity effect was reported by Nakayama (1988) using binocular disparity and occlusion in a timed recognition task. Although his effect occurred for raw response time, it is possible that RT slopes may be influenced in a similar way. An influence from proximity could explain the closer match between projective size predictions and findings when the target is short. Searching for the short target in the near location may be easy because it is closer and unexpectedly small in size. But when the long target is near, performance should be more difficult according to laws of projective geometry, yet easier according to proximity theory. These opposing effects could cancel out and account for why RT slopes do not differ when searching for a long target in near and far locations.

To test for an influence from apparent proximity, subjects could search for a target that is neutral to size information. One example might be to search for a gray item against black items in a texture gradient display. If both target and distractor items vary randomly in size, projective size influences would be controlled since item size is now irrelevant to the task.

Note that any influence from apparent proximity does not invalidate the findings of this thesis. Evidence of either apparent proximity or projective influences on search performance confirms that subjects perceive apparent depth from the texture gradient. Therefore, it is not critical that the depth effect be mediated by apparent proximity or assumptions of projective size. The only potential problem is that if both are operating, we can expect to see the influence only on short target trials. Since promising trends were revealed in absolute RT performance for both target size conditions, I will postpone a test for apparent proximity for follow-up research and pursue the test for assumptions of projective geometry using an alternative strategy.
Summary

Apparent depth influences on visual search were seen more reliably in absolute RTs than in RT slopes. Apparent depth acted only to slow performance in search conditions predicted to be more difficult (i.e., short target in the far location and long target in the near location) but did not speed up performance in the conditions predicted to be easier (i.e., short target in the near location and long target in the far location). Since performance remained within the attentive range for all locations (i.e., RT slopes > 10 msec/item) we know only that attention-based search (later vision) was influenced by apparent depth. These results suggest that the visual search test used in Experiment 2 may be inadequate for assessing the sensitivity of early vision to slant-from-texture. Modification of the test to permit such an assessment is described in Experiment 3.

Experiment 3:
Slant-from-texture influences in early vision

Although subjects used attention to search in all locations in the first two experiments, it does not preclude the possibility that preattentive processing can also extract depth information from texture gradients. The goal of Experiment 3 is to determine whether early vision is also influenced by slant-from-texture. Since search performance in the previous experiments was only impaired by the presence of slant, an alternative way to assess whether location effects persist in preattentive vision is to start out with a very easy task where search is performed in parallel. Such performance can be obtained by taking advantage of the findings that visual search performance can be adjusted along the preattentive-attentive dimension simply by changing the discriminability of the search items (Duncan & Humphreys, 1989; Treisman & Gormican, 1988). In the present study, sizes of the search items were adjusted for each subject. The aim was to have subjects performing parallel search on control trials. Any subsequent decrement in performance from searching for a short target in the far
location, or the long target in the near location would show that preattentive vision cannot ignore the apparent distance information implicit in the texture gradient.

Method

Subjects. Forty undergraduate students participated in two one-hour sessions, each consisting of six sets of 54 trials. Subject age ranged from 17 to 27 years; twenty-four subjects were female and sixteen were male. All subjects had no previous experience with visual search tasks, and had normal or corrected to normal vision.

Stimuli and Procedure. Visual search displays contained cylinder items on a textured background similar to the ones used in Experiment 2. Displays were modified in several ways to eliminate potentially confounding variables noted below. First, to control for luminance differences caused by the gradient of texture, the density of dark pixels making up the texture were systematically varied. When this procedure was completed, every 2.25 X 2.25 region contained 20% black pixels as is shown in Figure 8. Second, displays were presented in a square region, subtending 14° X 14°, to control for influences from different aspect ratios.

Each subjects' threshold for detecting the presence of a unique cylinder preattentively was assessed using a staircase method. The long target ranged in size from 1.40° x 1.10° to 1.72° x 1.42° across conditions in .04° (.35mm) increments; short distractors were held constant at 1.29° x .92° as is shown in Figure 9. Between four and six threshold tests were administered depending upon performance on preceding trials. Each threshold test contained 4 blocks of 20 trials. Display presentation and data collection were controlled by the VSearch program (Enns, Ochs, & Rensink, 1990).

-----Insert Figure 8-----

-----Insert Figure 9-----
Figure 8. Illustration of combined texture gradient background used in Experiment 3. Displays equate for luminance and overall aspect ratio.
Figure 9. Full set of search items used in staircase procedure to assess subject threshold for early vision. The long target ranged in size from $1.40^\circ \times 1.10^\circ$ to $1.72^\circ \times 1.42^\circ$ across conditions in $.04^\circ (0.35\text{mm})$ increments; short distractors were held constant at $1.29^\circ \times .92^\circ$. 
Changes to the stimuli and procedure for the primary visual search task were as follows: (1) The size of the long items, determined by the staircase tests, fell within the range of 1.49° x 1.17° to 1.65° x 1.33°. The size of the short items was held constant at 1.25° x .92°. (2) The horizontal orientation of the background texture was counterbalanced across subjects. Half of the subjects viewed slanted surfaces oriented so that the near region appeared on the left of the computer display, and the other half of the subjects viewed slanted surfaces with near on the right. (3) Search item orientation was also counterbalanced across subjects. Half of the subjects searched for items oriented horizontally on the screen. Items were oriented either left-to-right or right-to-left depending upon the direction of the surface in the slant conditions (i.e., when the surface was oriented with near on the left, and far on the right, cylinders were oriented with their bottom edge to the left, and their top edge to the right). These cylinders appeared perpendicular to the textured surface and to recede into the distance on top of the slanted surface. The remaining subjects searched for items oriented upright in the image. In the slant condition, these items appeared to rest on an image plane separate from the underlying slanted surface. (4) Display presentation and data collection for the search task were controlled by a Macintosh computer using the VScope Program (Enns & Rensink, 1992). (5) Feedback was presented at central fixation following subjects' response after each trial. Correct responses were indicated with a "+" and incorrect responses were indicated with a "-". (6) Performance asymmetries typically found in tests of visual search for item size (Treisman & Gormican, 1988) were also tested by exchanging the long and short target and distractor items.
Results

Threshold for preattentive vision.

All subjects met the criterion for preattentive search within six sets of 60 trials using the staircase procedure. The target and distractor size difference necessary for subjects to achieve preattentive performance, ranged from .25° to .40° of visual angle. Final RT slope values consistently fell within 4 and 10 msec/item on target present trials. Target absent trials were on average twice this magnitude.

Visual search on textured surfaces.

Mean correct search times across target present locations are reported in Table 3, along with the percentage of errors. The statistical analyses are based only on target present trials since absent trials are not differentiated by target location. Analyses comparing target presence and target absence trials are presented in Appendix A. Apparent depth effects are reported after a brief description of central region effects. Similar to Experiment 2, performance was always relatively better when the target was located in the middle location. This advantage in searching for the target in the central region is supported by significant comparisons of performance in middle and extreme locations (all t's, p < .05). Pairwise comparisons were based on Bonferroni t's that utilized MSe from the overall ANOVA. Further support for central region facilitation was found in the significant main effect for location that compared performance across the three locations (RT: \(F(2,70) = 116.8, p < .001\); errors: \(F(2,80) = 30.2, p < .001\)).

To assess for influences from apparent depth, data from the near and far locations were compared directly as shown in Figure 10. In the slant condition, search was on average 67 msec slower when the short target was far (left or right), and 27 msec slower when the long target was near (all t's, p < .05). Only when the short target was in the slanted condition was there a significant difference in percent errors across near and far
Table 3

Mean correct Reaction Time and Percentage of Errors in Experiment 3

Combined Gradient

<table>
<thead>
<tr>
<th>Display Size</th>
<th>Apparent Location</th>
<th>RT (msec)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long Target</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Far</td>
<td>754 (19.8)</td>
<td>17.3 (2.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>627 (13.3)</td>
<td>6.2 (1.2)</td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>614 (12.7)</td>
<td>3.0 (0.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>604 (12.5)</td>
<td>3.0 (0.9)</td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>658 (12.5)</td>
<td>5.9 (1.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>643 (13.5)</td>
<td>4.8 (1.1)</td>
</tr>
<tr>
<td>6</td>
<td>Far</td>
<td>758 (20.7)</td>
<td>6.7 (1.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>674 (14.3)</td>
<td>4.6 (1.2)</td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>626 (12.4)</td>
<td>3.5 (1.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>595 (10.9)</td>
<td>4.6 (1.2)</td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>684 (14.7)</td>
<td>3.8 (0.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>698 (19.8)</td>
<td>5.9 (1.1)</td>
</tr>
<tr>
<td>10</td>
<td>Far</td>
<td>814 (20.4)</td>
<td>9.2 (1.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>732 (15.6)</td>
<td>6.2 (1.2)</td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>673 (13.3)</td>
<td>1.9 (0.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>686 (15.0)</td>
<td>3.8 (1.0)</td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>781 (16.9)</td>
<td>9.4 (1.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>738 (18.7)</td>
<td>10.0 (1.7)</td>
</tr>
</tbody>
</table>

| Short Target |                   |           |            |
| 2            | Far               | 754 (19.8)| 17.3 (2.2) |
|              |                   | 627 (13.3)| 6.2 (1.2)  |
|              | Mid               | 614 (12.7)| 3.0 (0.9)  |
|              |                   | 604 (12.5)| 3.0 (0.9)  |
|              | Near              | 658 (12.5)| 5.9 (1.3)  |
|              |                   | 643 (13.5)| 4.8 (1.1)  |
| 6            | Far               | 758 (20.7)| 6.7 (1.4)  |
|              |                   | 674 (14.3)| 4.6 (1.2)  |
|              | Mid               | 626 (12.4)| 3.5 (1.0)  |
|              |                   | 595 (10.9)| 4.6 (1.2)  |
|              | Near              | 684 (14.7)| 3.8 (0.9)  |
|              |                   | 698 (19.8)| 5.9 (1.1)  |
| 10           | Far               | 814 (20.4)| 9.2 (1.6)  |
|              |                   | 732 (15.6)| 6.2 (1.2)  |
|              | Mid               | 673 (13.3)| 1.9 (0.7)  |
|              |                   | 686 (15.0)| 3.8 (1.0)  |
|              | Near              | 781 (16.9)| 9.4 (1.5)  |
|              |                   | 738 (18.7)| 10.0 (1.7) |

| Long Target |                   |           |            |
| 2            | Far               | 589 (14.5)| 6.5 (1.3)  |
|              |                   | 580 (10.8)| 7.5 (1.5)  |
|              | Mid               | 556 (12.2)| 3.8 (1.0)  |
|              |                   | 526 (8.4) | 2.2 (0.8)  |
|              | Near              | 613 (13.2)| 5.9 (1.3)  |
|              |                   | 558 (8.9) | 3.0 (0.9)  |
| 6            | Far               | 597 (11.5)| 5.1 (1.3)  |
|              |                   | 574 (10.0)| 3.2 (1.0)  |
|              | Mid               | 544 (9.7) | 2.4 (0.8)  |
|              |                   | 531 (12.9)| 2.2 (1.0)  |
|              | Near              | 622 (13.6)| 4.8 (1.1)  |
|              |                   | 578 (12.1)| 3.7 (1.2)  |
| 10           | Far               | 624 (12.2)| 6.7 (1.4)  |
|              |                   | 610 (11.8)| 6.2 (1.2)  |
|              | Mid               | 591 (11.1)| 5.9 (1.3)  |
|              |                   | 554 (10.8)| 4.0 (1.0)  |
|              | Near              | 654 (14.6)| 7.3 (1.4)  |
|              |                   | 582 (9.9) | 5.9 (1.2)  |
locations \(t(35), p < .05\). In the control condition, RT and percentage of errors were the same across left and right locations with the exception of the 10 item - short target condition; in this condition performance was opposite to the predicted apparent depth pattern \(t(35), p < .05\).

--- Insert Figure 10 -----

These trends are supported by a significant 3-way interaction between the factors of background, target size, and apparent depth location (i.e., Far and Near) on target present trials (RT: \(F(1,35) = 59.8, p < .001\); errors: \(F(1,40) = 8.1, p < .01\)). Difficulty in detecting the short target in the far location is further suggested by a significant target size by location interaction for RT (\(F(1,35) = 14.2; p < .001\)). These interactions also support the inverse relationship between short and long target in near and far slant conditions.

Further support for apparent depth having its influence in the slant condition was evident from the significant interaction between background and apparent depth (RT: \(F(1,35) = 9.4; p < .01\); errors: \(F(2,80) = 5.0; p < .05\)). This interaction reflects the high RTs and percentage of errors in the far region of the slanted background (mean RT: 689 msec; mean percent errors: 8.6 %) as contrasted with performance in the near region (mean RT: 668 msec; mean percent errors: 6.2%), or performance in either of the no slant locations (mean RT: 632 msec; mean percent errors: 5.6%). Finally, there was not a main effect for apparent location on RT performance \(F(1,35) = 2.4\), even though there were slightly more errors in the far region \(F(1,40) = 5.1; p < .05\).

**Display size and apparent depth.**

Evidence for an influence from projective assumptions was also expected in an apparent depth by display size interaction for the slanted conditions. In difficult search conditions (i.e., when the short target was in the far location), the influence of apparent depth was expected to increase with display size so that RT functions would appear to "fan out" across display sizes. Although this pattern was observed in Experiment 2, it did
Figure 10. Mean correct reaction time and percentage of errors for target present trials of Experiment 3. Two search items (short and long) were examined against three display sizes (two, six and ten) and two backgrounds (slant and control). Critical comparisons are between locations (far and near for the slant condition, and left and right for the control condition). Error bars are SEMs.
not occur here. The 4-way interaction between background, target size, apparent location, and display size was not significant ($F(2, 70) = 1.7$) nor do the means in Figure 10 indicate such an interaction.

Furthermore, if the 4-way interaction were significant, it would be quite a different pattern from the predicted one. Notice in Figure 10 that the RTs of the two and six item-far conditions appear much more inflated (relative to their near counterparts), than the ten item condition in short target trials. In support of this trend, pairwise $t$-tests revealed no difference between the two and six item displays across the two locations of the slanted background. If there was a significant 4-way interaction involving display size, it would indicate that apparent depth effects were strongest in short target conditions with the fewest number of search items.

For the long target conditions, display size interactions were completely absent. This is seen in the three parallel RT display size functions in all locations, as well as the nonsignificant pairwise comparisons across display sizes and apparent locations (all $t$'s, $p > .10$). The only significant display size effect was for the ten item trials where search was slower than two and six item trials for both the far and near locations ($t(70), p < .01$).

Subjects made the most errors when the short target was located in the far region of the two item display. A significant 3-way interaction between target size, background, and display size ($F(2, 80) = 4.5, p < .01$), as well as a significant 2-way interaction between display size and apparent location for percent errors ($F(4, 160) = 10.0, p < .001$), support the stronger apparent depth influence in small display size conditions.

The absence of the predicted 4-way interaction involving display size is evidenced further by separate simple effects on target size and background factor combinations. Only in an analysis of the short target - slanted background conditions was there a significant interaction between apparent location and display size ($F(2, 74) = 9.1, p < .01$). The corresponding trend showed apparent depth had its greatest
influence in the displays with the fewest items. The absence of the predicted apparent depth effects in the greater display size conditions is important since it shows that it is inappropriate to use RT slopes as a measure of the apparent depth influence.

Nevertheless, since RT slopes on target present trials did fall within the range of 4 and 11 ms/item (across locations) this ensures that early vision was used in all conditions.

Additional influences

The following analyses are included to assess secondary influences that could account for some of the noted apparent depth effects. These include influences from practice, top-to-bottom and left-to-right search biases, and potential biases introduced by item orientation.

Practice.

Main effects of trial block were not observed (RT: \( F(2,70) = 2.2 \); Error: \( F(2,80) = 1.7 \)). However, response time did improve over trials when there were many items in the display (block x display size -- RT: \( F(4,140) = 4.3, p < .001 \); errors: \( F(4, 160) = 0.5 \)). In addition, percent errors decreased with practice in the slant condition ( \( F(2,80) = 4.0, p < .05 \)), but there was no change in response time ( \( F(2,70) = 0.8 \)).

Top-bottom search biases.

Search performance across top, middle (horizontal meridian), and bottom locations are shown in Table 4. Overall, subjects were 37 msec faster with 2 % fewer errors when the target was located in the top region relative to the bottom region (RT: \( F(2,80) = 181.9, p < .001 \); errors: \( F(2,80) = 23.8, p < .001 \); \( t(80), p < .01 \)). Subjects were 32 msec faster and made 2% fewer errors when the target was located in the central region (\( t(80), p < .01 \)).

--- Insert Table 4 ---

After removing the middle location from the analysis, subjects were still significantly faster detecting the target when it was located in the top region (RT: \( F(1,40) \))
Table 4
Mean correct Reaction Time and Percentage of Errors in Experiment 3

Top - Bottom Analysis

<table>
<thead>
<tr>
<th>Display Size</th>
<th>Location</th>
<th>RT (msec)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slant</td>
<td>Control</td>
</tr>
<tr>
<td>Short Target</td>
<td></td>
<td>M (SEM)</td>
<td>M (SEM)</td>
</tr>
<tr>
<td>2</td>
<td>Top</td>
<td>663 (15.0)</td>
<td>638 (13.2)</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>646 (15.5)</td>
<td>591 (12.0)</td>
</tr>
<tr>
<td></td>
<td>Bottom</td>
<td>694 (13.2)</td>
<td>647 (11.9)</td>
</tr>
<tr>
<td>6</td>
<td>Top</td>
<td>684 (15.9)</td>
<td>662 (13.7)</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>655 (15.7)</td>
<td>647 (14.9)</td>
</tr>
<tr>
<td></td>
<td>Bottom</td>
<td>704 (14.4)</td>
<td>665 (13.8)</td>
</tr>
<tr>
<td>10</td>
<td>Top</td>
<td>720 (14.4)</td>
<td>722 (16.0)</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>708 (16.2)</td>
<td>676 (14.8)</td>
</tr>
<tr>
<td></td>
<td>Bottom</td>
<td>804 (18.6)</td>
<td>756 (14.6)</td>
</tr>
</tbody>
</table>

Long Target

| 2            | Top      | 582 (10.9)| 541 (7.3) | 3.5 (0.9) | 2.4 (0.9)  |
|              | Middle   | 563 (13.0)| 526 (7.5) | 4.1 (1.0) | 3.0 (0.9)  |
|              | Bottom   | 654 (16.6)| 588 (9.1) | 8.7 (1.5) | 7.3 (1.0)  |
| 6            | Top      | 615 (12.1)| 572 (12.1)| 6.2 (1.2) | 4.4 (1.1)  |
|              | Middle   | 559 (10.0)| 529 (7.5) | 2.2 (0.7) | 0.6 (0.4)  |
|              | Bottom   | 623 (12.8)| 597 (12.3)| 4.1 (1.1) | 4.3 (1.0)  |
| 10           | Top      | 626 (10.9)| 584 (9.9) | 6.0 (1.3) | 4.3 (1.1)  |
|              | Middle   | 587 (11.3)| 535 (9.1) | 2.7 (0.9) | 2.7 (0.8)  |
|              | Bottom   | 683 (14.4)| 629 (13.8)| 11.4 (1.8)| 9.2 (1.5)  |
Slant - from - texture

= 122.9, p < .001; errors: F(1,40) = 9.5; all t's, p < .01). Most of this difference is accounted for by the slant and long target conditions. Subjects were 45 msec faster to detect the target at the top of the slanted condition, as compared to 26 msec in the control condition (t(40), p < .001). Subjects were also 79 msec faster to detect the long target in the top region, as compared to 30 msec faster for the short target (t(40), p < .001).

Supporting statistics include a significant background by location interaction (RT: E(1,40) = 8.1, p < .01; errors: E(1, 40) = .01), and a target size by location interaction (RT: E(1, 40) = 4.2, p < .05; errors: E(1, 40) = 2.0). The possibility that these difference across target size and location conditions may be related to a height-in-plane depth interpretation is considered in the discussion.

Location also interacted with display size (RT: E(2,80) = 14.6, p < .001; errors: E(2, 80) = 9.9) and together these interacted with target size (RT: E(2,80) = 3.7, p < .05; Error: E(2, 80) = 0.3). With the exception of the 6 item - short target condition, there was a significant difference in performance between top and bottom locations when collapsed across background displays (all t's, p < .05). The greatest difference of 60 msec occurred in the two item - long target trials and also in the ten item - short target trials.

**Left - right search bias.**

Subjects were 19 msec faster in detecting the target when it was located on the left side of the control display. This left side advantage, although small, was supported by a significant interaction between background and location for RT (E(1,39) = 11.0, p > .01), but not errors (E(1,40) = .8). This trend occurred only in conditions with six (t(40), p < .01) or ten items (t(40), p =.09) in the display. When only two items were present, subjects were better at detecting the target on the right side of the display. These distinct trends across display size and location are supported by a significant interaction between location and display size for both RTs (E(2, 78) = 10.2, p > .001);
and errors ($F(2,80) = 3.8, p < .05$). Perhaps these trends reflect the necessity of multiple items for a lateral search bias to emerge. One final interaction between target size and location reached significance, ($F(1,40)=7.1, p < .01$), indicating that subjects made fewer errors when the short target was located on the left side of the display screen. Note that these search biases cannot account for apparent depth effects.

**Item Orientation.**

Figure 11 presents performance across vertical and horizontal search item conditions for three factors -- target size, background, and apparent location. Search on the slanted surface was slower for horizontal items than vertical items with greater differences across apparent location occurring on long target trials ($t (34), p < .01$), and horizontal orientations ($t (34), p < .05$). For the short target slant conditions there were significant differences between apparent location for both the vertical and horizontal conditions ($t (34), p < .01$). The effect of item orientation on the long target trials, but not short trials, were supported by an interaction between item orientation, background, and apparent location for RT performance ($F(1,34) = 4.6, p < .05$).

--- Insert Figure 11 ----

Item orientation also shows consistent influences across the different display sizes as reflected in the absence of a further interaction with display size (RT: $F(2,68) = .18$; Error: $F(2,78) = .46$). However, display size did interact with item orientation on its own, (RT: $F(2,68) = 7.6, p < .001$; Error: $F(2,78) = 4.8, p < .01$), and these two factors interacted with target size for RT, ($F(2, 68) = 3.2, p < .05$). These interactions reflect an undifferentiated increase in RT with increases in display size for the horizontal conditions, one that was strongest on short target trials. Such a trend is indicative of a manipulation that makes search more difficult.

The mixed-design ANOVA revealed a marginally significant interaction between item orientation and the remaining factors on percent errors ($F(1, 39) = 3.3, p = .07$),
Figure 11. Mean correct reaction time and percentage of errors for target present trials of Experiment 3 separated across item orientation. Filled circles and bars represent data from horizontal item trials; open circles and bars represent vertical item trials. Error bars are SEMs.
but not RT ($F(1, 34) = 0.4$). This interaction shows that error performance on horizontal conditions were most consistent with the predicted apparent depth influence in that they were greatest when the short target was far. Pairwise comparisons support this trend with the only significant difference between the far and near locations occurring on the short target slant conditions that contain horizontal items ($t (34), p < .01$).

**Size - depth consistency.**

To further investigate the influence of item orientation on apparent depth, target size and apparent depth were combined to form a factor representing *size - depth consistency*. The *consistent* condition represents search for the short target in the far location, or the long target in the near location, and the *inconsistent* condition represents search for the short target when it was near, or the long target when it was far. Notice that this factor collapses means across difficult search conditions so that apparent depth influences now share the same prediction -- search should be more difficult in the consistent conditions and easier in the inconsistent conditions. Instead of having to rely on the interaction between target size and apparent location in the slant condition to demonstrate apparent depth influences, only a main effect for size-depth consistency is required.

Interactions involving item orientation are reported since this is the variable of interest. Consistent trials were substantially slower than inconsistent trials when the items were oriented horizontally and subjects were searching for the long target (item orientation, target size, background by size-depth (RT: $F(1,34) = 4.6, p < .05$; errors: $F(1,39) = 2.4$). When subjects searched for the short target, this difference emerged only in error performance ($t (34), p < .05$). Although the corresponding 4-way interaction noted above was not significant for error performance, there was a marginally significant interaction between item orientation, background and size-depth
consistency ($F(1, 34) = 3.3, p < .07$) reflecting that horizontally oriented items produced the greatest difference across consistent and inconsistent slant conditions.

**Discussion**

The visual search study presented here shows that early vision is influenced by slant-from-texture. The main evidence was a shift in performance from a fast to slow search in conditions where projected size rendered the size discrimination task more difficult (i.e., when the short target was far or the long target was near). Further evidence that these are early vision effects is reflected in the consistently shallow RT slopes across all conditions. The implication is that subjective depth perception does relate to the underlying processes governing search, and that attention is not necessary for depth processing.

The larger apparent depth effect in horizontal trials (seen in RT data for long targets and percent error data for short targets), further supports the thesis that early vision is sensitive to apparent depth. Search items in this condition appeared to be attached to the underlying slanted surface. Recall the description of the appearance of the items in the method -- these cylinders appeared perpendicular to the textured surface and thus receded into the distance with the slanted surface.

**Item grouping?**

RT functions were expected to "fan out" across display sizes as usually occurs in a difficult search task. Instead, apparent depth effects either did not differ across display sizes, as in the long target trials, or were strongest in short target conditions with the fewest number of search items. The small display size effect in the long target condition can be explained by an overall reduction in reaction times, but the reversed display size effect in the short target condition must be explained by an alternative influence -- perhaps item grouping. Manipulating the appearance of the underlying surface slant can provide the conditions for item grouping. Items that appear in a
separate plane from the apparent depth are easier to ignore. As the number of items increase in the display, the apparent plane formed by the items is strengthened. This possibility of an item grouping influence is further supported by orientation influences, where weaker apparent depth influences were found in conditions with vertically oriented items. The inconsistency between the flat plane of the items and the slant of the background on these trials may have reduced the effect of the texture on search performance.

Search biases.

Results showed that subjects were faster to detect the target when it was at the top than when it was at the bottom of the display. A simple advantage in detecting the target in the top region could be explained by a systematic search bias, perhaps adopted from reading experience. However, this top advantage was much more pronounced for long target trials than short target trials. Such performance differences across target size may reflect height-in-plane influences noted in the previous experiment, whereby the top region may be interpreted, to some degree, as further away. A height-in-plane interpretation should manifest itself in harder search in the top region for the short target, and easier search in the top region for the long target. Height-in-plane effects in the short target trials may have been reduced somewhat by competing top-to-bottom influences where search in general is easier at the top of the display. Alternatively, height-in-plane interpretations may be more pronounced in later stages of visual processing. This interaction between item orientation and background is discussed further in the presentation of visual search performance in the isolated gradient conditions of Chapter 4.

Implications.

Evidence for sensitivity to apparent depth in early vision has interesting implications for the representations that are formed there. One in particular concerns
whether these representations correspond to the retinal image, or whether they have been transformed through a mechanism that adjusts for projected size to reflect properties of external objects. By and large, researchers who work with the visual search paradigm adhere to an analytic view of perception, arguing that preattentive representations resemble the retinal image. Their premise is that the experience of complex wholes is built from elementary properties, and that perceptual processing first involves some decomposition of the visual input into separate dimensions or components. Considerable psychophysical (Beck, 1966, 1967; Julesz & Bergen, 1983; Treisman & Gelade, 1980) and physiological support exists for this view (i.e., Hubel & Wiesel, 1977; Barlow & Levick, 1965; Blakemore, 1975; Campbell & Robson, 1968; Cowey, 1979; Zeki, 1978; 1981). Its proponents typically regard more primitive features (i.e., size, orientation, and color) as the basic elements of vision and cite evidence showing that such features are independently coded and stable across different contexts.

The opposing view that global attributes (i.e., surface properties such as overall shape or orientation) are detected before features (i.e., size) is supported by demonstrations of global precedence (Navon, 1977); object identification preceding brightness perception (Coren, & Komoda, 1973); and improved object-part identification in the context of three-dimensional objects (Enns & Gilani, 1988; Lanze, Maguire, & Weisstein, 1985; Lanze, Weisstein, & Harris, 1982; Weisstein & Harris, 1974). Texture segmentation and visual search evidence in favor of this view include influences of global grouping on visual search (Banks & Prinzmetal, 1976); irrelevant variation of features slowing texture segmentation (Callaghan, 1989), and redundant variation among features speeding up segmentation (Callaghan, 1989; Callaghan, Lasaga, & Garner, 1986). There is also evidence showing context effects in visual search, as well as automatic depth detection (Aks & Enns, 1992; Enns & Rensink,
1990a; 1990b; Kleffner & Ramachandran, 1992; Nakayama & Silverman, 1986; and Ramachandran & Plummer, 1989). These studies, together with the present findings, show that an apparent depth context can influence the earliest stages of perception in predictable ways.
Chapter 4: An examination of the texture gradient dimensions

The switch from fast to slow search in conditions where the short target is *far* and the long target is *near* shows that early vision is sensitive to the apparent depth in the texture gradient. One way to reconcile these data with current theory on early vision is to redefine visual information in terms of environmental coordinates rather than the 2-D retinal coordinates. Experiment 4 investigated the question of which information contained within the texture gradient contributes most to the 3-D interpretation of early vision.

**Monocular Depth Cues**

Let us take a brief digression to further describe available depth information that is related to the texture gradient. After all, it is a logical precursor to understanding what 3-D information is used by the human visual system. Only monocular depth cues are described since texture gradients fall within this category of depth cues.

*Occlusion* provides explicit information about front-to-back object relations in our frontal plane. Stimulus components that conceal other components are perceived as lying in front.

*Shadows* provide information about convexity and concavity as well as information about object position (Cavanagh & Leclerc, 1989). This shape and depth information can be used if one assumes that light comes from above.

*Familiarity* enables us to maintain size-constancy by comparing stimuli in our visual field with stimuli in our memory. Consequently, we can perceive changes in distances when the proximal representation of an object changes its size.

*Gradients* of motion, shading, blur, and texture, all can be viewed as information that systematically grades, or changes within the visual field as a direct function of changes in distance and shape. For instance, motion parallax, informs us about distance because close objects are seen as moving faster than further objects.
Shading gradients provide depth information by signifying smooth depth changes of convexity or concavity (Mingolla & Todd, 1986; Ramachandran, 1988; Todd & Mingolla, 1983). Gradients of blurring can originate from either our own physical limitations or from environmental interference. The former can refer to retinal blurring when we are unable to resolve fine details of receding and hence, shrinking images. The latter can refer to interference from haze or dust. Because both forms of blurring systematically grade with receding objects, it is reasonable to classify these as gradient-like cues. Finally, texture gradients, being dominated by changing retinal size information, incorporate a number of the properties noted above.

Additional important characteristics of all these monocular depth cues is that they interact with one another (i.e., familiarity can override size information), contribute to one another (i.e., element density gradients frequently are accompanied by shading gradients) and contain information that can be treated as separate depth cues (i.e., perspective and compression). These same properties apply to all of the dimensions that make up the texture gradient cue. For now, we will consider this particular gradient cue as an independent source of depth information that can contain within itself an entire subset of depth cues.

Decomposing the texture gradient.

Since gradients can be regarded as sources of information that grade with the visual angle, this implies that the dimensions can be logically separated. Using Cutting and Millard’s (1984) terminology, gradient information available in the environment includes perspective, compression, and density. Figure 12 shows perspective and compression in their separated and combined forms.

--- Insert Figure 12 ---

The perspective gradient refers to one type of convergence information that is defined as the change in width of the textural element. It is also referred to here as the element
Figure 12a. Reduced illustrations of visual-search displays from Experiment 4 containing vertical items. Two search items (long and short) were examined against six search backgrounds -- combined, perspective compression and corresponding controls.
Figure 12b. Reduced illustrations of visual-search displays from Experiment 4 containing horizontal items. Two search items (long and short) were examined against six search backgrounds -- combined, perspective compression and corresponding controls.
width / element height, with height held constant across the gradient pattern.

Compression gradient is synonymous with the artistic convention of foreshortening, and is represented by the change in element height / element width. Here, width is held constant across all the texture elements. (Notice that the gradients in this figure are rotated sideways. Although width still refers to the horizontal dimension of the elements making up the gradient, it corresponds to the vertical dimension on the page. The opposite is true of height.) Density gradients refer to the increasing number of texture elements per unit of visual angle. The density gradient can be defined in terms of a perspective or a compression gradient. A density gradient in the width of the texture elements can be classified as a perspective gradient, and a density gradient in the height can be classified as a compression gradient. For purposes of clarity and consistency, the terms perspective, compression, and density will be used throughout the remainder of this thesis, even when describing research that uses interchangeable terms such as aspect ratio, form ratio, area, scale, height, and size.

Relative contribution of the texture dimensions to subjective perception.

The most common methods used to test the relative importance of each cue to the perception of depth include isolating individual depth cues, pitting one against the other, and recording judgments about slant or depth (Braunstein, 1976). Typical procedures involve asking subjects about the existence, the direction (towards or away), and the extent of the phenomenal depth impression that is evoked by the different information contained within texture gradients. A series of studies using these techniques found that gradients of perspective and compression have a greater impact than do gradients of density on judgments of slant (Braunstein, 1976; Braunstein and Payne, 1969). Notice that in the present experiment, with the textured surface consisting of lines, density is inseparable from perspective or compression.
Flat slanted surfaces can be regarded as only a small subset of surface properties found in the natural environment; accordingly, the study of the perception of curved surfaces has also been investigated. In keeping with this shift towards greater ecological relevance, Cutting and Millard (1984), employed a preference and dissimilarity paradigm to determine the relative importance of different depth cues in 3-D judgments of flatness and curvature. In the preference procedure, subjects selected the texture gradient pattern which appeared flat and the pattern which appeared curved. In the dissimilarity measure subjects rated, on a scale from 1 to 5, the extent to which various texture gradients appeared either flat or curved. Cutting and Millard confirmed that different texture information is used in the perception of flat versus curved surfaces. In the perception of flat surfaces, the variance explained by each dimension of the texture gradient was: 65% by the perspective gradient, 28% by the density gradient, and only 6% by the compression gradient. This contrasted with perceived curvature, with 96% of the variance accounted for by the compression gradient, and less than 2% by the other two gradients.

Similar findings were reported in subsequent research that tested the perception of textured images depicting spherical objects (Todd & Akerstrom; 1987). Texture gradient dimensions were isolated by randomizing variations in size, shape, and orientation, as well as manipulating the horizontal viewing angle which is known to influence the perspective gradient (i.e., width variation). The results showed that systematic variations in compression, together with the appropriate orientation positions for individual elements, contribute most to accurate curvature judgments. The latter orientation effects extend the findings of Cutting and Millard (1984), in showing the insufficiency of local texture information and the importance of a global gradient in eliciting accurate impressions of curvature (but see, Stevens; 1981).
The literature on texture gradient perception shows that different information is used in perceiving flat and curved surfaces and that this information provides differently weighted contributions to our final impression of depth. Perspective information is used mostly for the perception of flat slanted surfaces, followed by density and then compression. For the perception of curvature, compression information is the most informative with slight contributions from perspective and density.

**Experiment 4: Influences from dimensions of the texture gradient on visual search**

Having established a tool for reliably measuring the slant-from-texture influence on visual processing, we now can assess the relative influences of the texture gradient dimension. Given the evidence for separate contributions from the dimensions in subjective perception we might also expect separate contributions in visual search. However, this prediction must be made cautiously, given the dissociation between subjective perception and early perception discussed in Chapter 1.

In Experiment 4, the perspective and compression dimensions of the texture gradient were separated across visual search conditions, as shown in Figure 12. Because these cues provide different spatial information (i.e. Cutting & Millard, 1984; Stevens, 1981; 1984), we may consequently have developed corresponding functions in our visual systems. Stevens (1981) has demonstrated that these dimensions specify distinct properties of the environment -- perspective informs us about distance, and compression informs us about orientation (Cutting, 1984; Stevens, 1981; 1984; Witkin, 1981). Perspective (i.e., the characteristic dimension) is the only dimension affected solely by distance. All other dimensions of the projected texture are affected by distance, orientation, and object height. Perspective is scaled only in width while compression is scaled in height and width. Given that perspective is the only pure
measure of texture distance, and that compression is the dimension most foreshortened by changes in slant, these two dimensions together are likely to be among the most reliable means for extracting information about distance and slant.

Stevens (1984) further speculates that perspective and compression may map onto depth and orientation representations similar to those proposed by Marr (1982) and Ullman (1984). According to this view, the distance information obtained from perspective can be coded independently of the orientation information derived from compression. The advantage of independent coding is evident in the many situations where only one form of this information is available or recoverable.

Support for such a distinction is available in the psychophysical literature. Cutting and Millard (1984), for example, found a performance asymmetry for these two dimensions in judgments of flat (i.e., maximal distant information) and curved (i.e., maximal orientation information) representations -- the perspective gradient was used most in judging the slant of flat surfaces and the compression gradient was used most in interpreting surface curvature.

Before assessing the separability of these texture gradient dimensions in visual performance, Experiment 4 first examines their relative impact on early visual processing. Experiment 5 takes the further step in assessing the independence of these texture gradient dimensions in early visual processing.

**Method**

Subjects, stimuli, and procedures were the same as those used in Experiment 3, except for the addition of displays that isolate the dimensions of the texture gradient. The six conditions compared in this experiment included (1) combined gradients, (2) vertical and horizontal line grid (control for the combined gradients), (3) perspective gradient, (4) vertical line grid (control for the perspective gradient), (5) compression
gradient, and (6) horizontal line grid (control for the compression gradient). These are shown in reduced form in Figure 12. The perspective and compression gradients are used to test the apparent depth influence of the separate dimensions on the search task relative to the combined dimension condition. The corresponding control grids test for vertical or horizontal background influences on search and also provide baselines with which to compare apparent depth influences.

As in Experiment 3, subjects viewed search items that were either vertical or horizontal in their orientation. Testing these two orientations is especially important in this experiment because it is possible that search items of different orientations may interact differently with the dimensions of the texture gradient.

Results: Perspective & Compression Combined

Target present data from the combined perspective and compression gradient condition in Experiment 3 were collapsed across background conditions to further isolate influences from early assumptions of projective size. The difference for search in the slant and control backgrounds was used as a measure of the apparent depth influence. These are shown in Figure 13 along with the analyses of the individual dimensions described below. Apparent location and target size were combined into a single factor of size-depth consistency for the ANOVA, as in Experiment 3.

Consistent with the previously noted trends, there was a greater difference across backgrounds in the size-depth consistent condition relative to the inconsistent condition (RT: $F(1,34) = 44.3, p < .001$; errors $F(1,39) = 8.8, p < .001$). Short target trials showed a larger difference than long target trials (RT: $F(1,39) = 4.2, p < .05$; errors: $F(1,39) = 6.6, p < .01$), and of the display size conditions, the two item trials showed the greatest difference (RT: $F(2, 68) = 3.3, p < .001$; errors $F(2, 78) = .40$). Errors
Figure 13a. Mean reaction time difference and percentage error difference across slant and control background for the perspective, compression and combined gradients in Experiment 4. Short targets are presented against three display sizes (two, six and ten) and two backgrounds (slant and control). Critical comparisons are between size-depth consistent (short-far) and inconsistent conditions (short-near). Error bars are SEMs.
Figure 13b. Mean reaction time difference and percentage error difference across slant and control background for the perspective, compression and combined gradients in Experiment 4. Long targets are presented against three display sizes (two, six and ten) and two backgrounds (slant and control). Critical comparisons are between size - depth consistent (long - near) and inconsistent conditions (long - far). Error bars are SEMs.
occurred most frequently in the two item-short target conditions (target size by display size, $F(2, 78) = 3.5$, $p < .001$).

**Item orientation.**

Figure 14 shows performance separated across the vertical and horizontal item trials and collapsed across display size. Only long target trials showed a significant effect from item orientation -- the horizontal items produced reliable apparent depth effects, unlike vertical item conditions (target size, size-depth consistency by item orientation -- $F(1,34) = 4.6$, $p < .05$; errors $F(1,39) = 1.6$). The difference across consistent and inconsistent trials was highly significant for all target size and item orientation trials ($t(34), p < .05$) except the long target-vertical conditions. Pairwise comparisons (based on Bonferonni $t$'s) of consistent versus inconsistent trials across display sizes were significant in all but three vertical item conditions: short target-ten item trials, and long target-two and six item trials ($t(34), p < .05$). Only marginal significance was obtained for the error data with a similar trend of horizontal items producing greater differences across consistent and inconsistent trials (size-depth consistency by item orientation -- $F(1,39) = 3.2$, $p = .08$).

------ Insert Figure 14 ------

**Results: Perspective**

Visual search data for the target present trials of the perspective gradient condition are shown in Table 5. Search for the long target was 17 msec faster when it was far ($t(70), p < .01$), and RTs and errors did not significantly differ across locations on short target and control trials. Search was also 53 msec faster when the target was located in the middle region ($F(2,70) = 78.9$, $p < .001$; $F(2,76) = 31.2$, $p < .001$). An additional ANOVA that omitted central region data was performed to remove influences from the overall faster performance in that region. The RT
Figure 14a. Short target trial mean reaction time difference and percent error difference across slant and control backgrounds in Experiment 4 separated by item orientation for each of three background combinations (perspective, compression, combined conditions and corresponding controls). Filled circles and bars represent data from horizontal item trials; open circles and bars represent vertical item trials. Error bars are SEMs.
Figure 14b. Long target trial mean reaction time difference and percent error difference across slant and control backgrounds in Experiment 4 separated by item orientation for each of three background combinations (perspective, compression, combined conditions and corresponding controls). Filled circles and bars represent data from horizontal item trials; open circles and bars represent vertical item trials. Error bars are SEMs.
difference across far and near locations in the long target condition was confirmed by a
significant interaction between the factors of background, target size, and apparent location
(RT: F(1,37) = 4.4, p < .05; errors: F(1,38) = 0.2). Further support for apparent depth
effects in the long target trials was found in an analysis based on background differencing
described below.

Display size and apparent depth.

Although apparent depth appears in Table 5 to have a greater impact on conditions
with fewer display items in the short target condition, this trend was not supported by the
predicted four-way interaction between background, apparent location, target size, and
display size (RT: F(2,74) = .37; errors: F(2,76) = 1.0); nor were the corresponding
pairwise comparisons significant. Two interactions involving display size were significant
but are of secondary importance. One showed a strong positive linear relation between RT
and display size for short target trials (display size X target size -- RT: F(2,74) = 7.5, p <
.001; errors: F(2,76) = 12.3, p < .01). The second showed that the difficulty of the ten
item conditions was most pronounced in the short target - slant condition, and the long
target - control condition (display size X target size X background -- RTs (F(2, 74) = 3.2,
p < .05; errors: (F(2,76) = 1.7).

--- Insert Table 5 -----

Performance difference across backgrounds.

Differences in performance across the slant and no slant backgrounds were
examined in the same manner as in the combined gradient condition. This included
combining target size and apparent location as a size-depth factor. The differences across
the perspective background and the horizontal control grid are shown in Figure 13.

In accordance with apparent depth predictions, the difference across backgrounds
in the consistent conditions (19 msec) was highly significantly relative to the null
Table 5

Mean correct Reaction Time and Percentage of Errors in Experiment 4

<table>
<thead>
<tr>
<th>Display Size</th>
<th>Apparent Location</th>
<th>RT (msec)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slant</td>
<td>Control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M (SEM)</td>
<td>M (SEM)</td>
</tr>
<tr>
<td>2</td>
<td>Far</td>
<td>659 (16.8)</td>
<td>617 (14.7)</td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>594 (12.9)</td>
<td>566 (9.9)</td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>646 (15.3)</td>
<td>619 (10.6)</td>
</tr>
<tr>
<td>6</td>
<td>Far</td>
<td>662 (14.8)</td>
<td>642 (13.1)</td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>590 (11.9)</td>
<td>596 (12.2)</td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>657 (18.2)</td>
<td>665 (14.5)</td>
</tr>
<tr>
<td>10</td>
<td>Far</td>
<td>707 (17.2)</td>
<td>689 (13.9)</td>
</tr>
<tr>
<td></td>
<td>Mid</td>
<td>657 (12.7)</td>
<td>630 (11.0)</td>
</tr>
<tr>
<td></td>
<td>Near</td>
<td>731 (19.8)</td>
<td>701 (16.3)</td>
</tr>
</tbody>
</table>

Long Target

| 2            | Far               | 563 (10.8)| 567 (9.5) | 5.4 (1.1) | 6.2 (1.3) |
|              | Mid               | 524 (7.7) | 520 (8.6) | 1.1 (0.6) | 2.0 (0.7) |
|              | Near              | 577 (11.6)| 574 (12.9)| 5.4 (1.2) | 3.9 (1.0) |
| 6            | Far               | 573 (12.2)| 575 (10.6)| 5.1 (1.2) | 5.1 (1.2) |
|              | Mid               | 539 (9.7) | 530 (10.8)| 3.4 (0.9) | 2.8 (0.9) |
|              | Near              | 595 (12.9)| 567 (9.2) | 3.7 (1.1) | 4.2 (1.0) |
| 10           | Far               | 611 (14.4)| 608 (14.1)| 7.4 (1.4) | 5.7 (1.2) |
|              | Mid               | 577 (12.0)| 567 (11.4)| 4.2 (1.1) | 3.7 (1.1) |
|              | Near              | 625 (15.0)| 593 (10.6)| 4.8 (1.1) | 4.0 (1.1) |
difference in the inconsistent conditions. This trend is supported by a main effect for size-depth consistency (RT: $F(1, 35) = 8.2, p < .01$; errors ($F(1, 38) = 0.2$). When separated across target size and display size conditions, only the short target - two item and the long target - ten item displays showed a significant RT difference for the consistent condition over the inconsistent condition ($t(38), p < .05$). When collapsed across display size only the long target showed a significant difference across consistent and inconsistent trials ($t(38), p < .05$). There was also a significant interaction between target size and display size (RT: $F(2, 70) = 3.7, p < .05$; errors ($F(2,76) = 0.2$), reflecting the high RT difference in the two- and ten-item short target conditions.

These results show that perspective does produce a significant difference across size-depth consistent and inconsistent conditions thus indicating an apparent depth influence in early visual processing. These effects are reliable across long target trials but are limited to small display sizes in the short target trials. The following sections report analyses of influences from item and background orientation. These tests are of secondary interest but are included to check for potentially confounding influences on apparent depth performance.

**Item orientation.**

The middle panel of Figure 14 shows performance in the perspective conditions separated across the vertical and horizontal item trials. From these trends it is clear that item orientation accounts for some of the performance trends noted above. The strongest apparent depth influence occurred in short target - vertical item trials, and in long target - horizontal trials ($t(35), p < .05$). A mixed-design ANOVA, with item orientation as a between factor, supported this asymmetry with a significant interaction between size-depth consistency, target size, and item orientation (RT: $F(1, 35) = 6.4, p < .05$; errors:
Additional location influences.

Overall, subjects were 23 msec faster when the target was in the top third as compared to the bottom third of the display (RT: $F(1, 40) = 44.0, p < .001$; errors: $F(1, 40) = 7.1, p < .05$). This difference is accounted for primarily by the long target conditions as supported by a significant target size X location interaction (RT: $F(1, 40) = 31.5, p < .001$; errors: $F(1, 38) = 3.0$). There were also no significant differences across left and right locations in the control condition. However, subjects were marginally faster (15 msec) detecting the target on the left in the slant condition (background by location -- RT ($F(1, 39) = 3.8, p = .06$), errors ($F(1, 38) = 1.2$).

Summary of perspective effects

Apparent depth effects were found in trials where the perspective gradient was present in the background. This effect was fairly robust across target size and display size conditions especially for long target trials and short target trials containing the fewest number of search items. Secondary influences from top-bottom or left-right search biases could not account for these trends. Stronger apparent depth effects occurred when items were horizontal on long target trials, and vertical on short target trials. This asymmetry across target size may be due to the local contrast of the items' orientation relative to the background (e.g., Gillam, 1973).

Results: Compression

Visual search data for target present trials of the compression gradient condition and its control are shown in Table 6. The analysis of the far, middle, and near location showed only a significant location effect for the mid-region which was on average 52 msec faster than either of the extreme locations (RT: $F(2, 78) = 103.7, p < .001$;
Differences in performance across near and far locations only emerged in particular display size conditions as described below.

When central region data were removed from the analysis, the 4-way interaction between background, target size, apparent depth, and display size was significant (RT: $F(2, 72) = 3.5, p < .05$; errors: $F(2, 78) = .08$). Search was slower when the short target was far in the two item displays, but faster in the ten item conditions ($t(72), p < .05$). The only other significant interaction was between apparent location and display size for errors ($F(2, 78) = 3.5, p < .05$), reflecting the poorest performance in the far location of the two item trials. In long target conditions there were no significant differences for far and near locations. However, significant effects do emerge in analyses described below that isolate background differences and size-depth consistency.

Performance difference across backgrounds.

A direct comparison of the difference in performance across backgrounds is shown in the right panel of Figure 13. In the analysis of size-depth consistency, the main effect for size-depth consistency did not reach significance (RT: $F(1, 39) = 1.2$; errors ($F(1, 38) = 1.2$) indicating that the consistent condition did not significantly differ from the inconsistent condition across all target size and display size conditions. However, a simple effects test isolating the long target condition did show a significant main effect for size-depth consistency (RT: $F(1, 39) = 4.8, p < .05$). Here subjects were slower in the consistent condition relative to the inconsistent condition ($t(39), p < .05$).

In the ANOVA based on short and long target sizes, size-depth consistency interacted with display size (RT: $F(2, 76) = 3.2, p < .05$; errors ($F(2, 78) = .12$), and together these interacted with target size (RT: $F(2, 76) = 3.3, p < .05$; errors ($F(2, 78) = .
## Table 6

**Mean correct Reaction Time and Percentage of Errors in Experiment 4**

**Compression Gradient**

<table>
<thead>
<tr>
<th>Display Size</th>
<th>Apparent Location</th>
<th>RT (msec)</th>
<th>Errors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slant</td>
<td>Control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M (SEM)</td>
<td>M (SEM)</td>
</tr>
</tbody>
</table>

### Short Target

| 2 | Far  | 685 (18.5) | 640 (18.4) | 6.6 (1.3) | 5.5 (1.2) |
|   | Mid  | 582 (12.5) | 585 (13.7) | 1.7 (0.7) | 3.9 (1.0) |
|   | Near | 635 (13.4) | 652 (14.5) | 3.3 (0.9) | 4.4 (1.0) |
| 6 | Far  | 663 (14.2) | 662 (17.4) | 5.0 (1.2) | 6.1 (1.4) |
|   | Mid  | 600 (10.8) | 593 (11.8) | 3.0 (0.9) | 2.5 (0.9) |
|   | Near | 661 (15.7) | 643 (14.9) | 3.3 (0.9) | 4.1 (1.0) |
| 10 | Far | 710 (16.2) | 708 (15.9) | 11.6 (1.7) | 8.6 (1.5) |
|    | Mid | 649 (13.7) | 648 (12.7) | 4.1 (1.0) | 4.4 (1.0) |
|    | Near | 746 (20.7) | 702 (17.0) | 10.5 (1.6) | 7.5 (1.4) |

### Long Target

| 2 | Far  | 583 (9.6) | 574 (10.5) | 5.2 (1.1) | 3.9 (1.1) |
|   | Mid  | 550 (11.8) | 534 (10.7) | 4.1 (1.1) | 3.6 (1.0) |
|   | Near | 594 (11.5) | 563 (11.5) | 6.6 (1.3) | 4.1 (1.1) |
| 6 | Far  | 596 (14.2) | 574 (9.8) | 3.3 (0.9) | 3.6 (1.1) |
|   | Mid  | 544 (9.2) | 543 (11.1) | 3.3 (0.9) | 1.7 (0.7) |
|   | Near | 601 (16.3) | 559 (8.6) | 5.2 (1.2) | 4.1 (1.1) |
| 10 | Far | 611 (12.1) | 608 (11.3) | 5.8 (1.2) | 6.1 (1.3) |
|    | Mid | 580 (12.9) | 568 (10.5) | 3.0 (0.9) | 4.4 (1.2) |
|    | Near | 634 (15.5) | 599 (11.6) | 6.9 (1.4) | 4.4 (1.1) |
Consistent trials were slower than inconsistent trials for two item displays collapsed across target size ($t(78)$, $p < .05$) in short target trials ($t(78)$, $p < .05$) and marginally in long target trials ($t(78)$, RT: $p = .10$; error: $p = .07$). When ten items were present on short target trials, the trend was reversed -- subjects took longer to detect the target on inconsistent trials ($t(78)$, $p < .05$). This pattern of performance may be a consequence of the grouping effect described in the Discussion, as well as the orientation influence described below. The following tests for potentially confounding influences are of secondary importance.

**Item orientation.**

The right panel of Figure 14 shows RT differences separated across vertical and horizontal item orientations in the compression conditions. Horizontal items produced slightly slower search as reflected in the marginally significant main effect for item orientation $F(1,35) = 3.2$, $p = .08$). The poorer performance was most pronounced in size-depth consistent conditions for errors ($F(1, 38) = 8.0$, $p < .001$). Pairwise comparisons of consistent versus inconsistent conditions were significant only in long target - horizontal trials ($t(38)$, $p < .05$).

**Additional influences.**

Subjects were 31 msec faster when the target was in the top third relative to the bottom third of the display (RT: $F(1,39) = 53.1$, $p < .001$; errors: $F(1,38) = 21.7$, $p < .001$; $t's$, $p < .001$). Subjects were also faster with the target located on the left under two conditions -- when ten items were present in the display (location X display size -- RT: $F(2,78) = 13.5$, $p < .001$; errors: $F(2,78) = 16.7$, $p < .001$) and when the target was short (location X target size -- $F(1,37) = 15.2$, $p < .001$).
Summary of compression effects.

Apparent depth effects were found when compression was present on its own in the background of the display. This effect was most most reliable in the long target trials. For short target trials, the effect was present only when two items were in the display. Stronger apparent depth effects also occurred when items were horizontal. Secondary influences from top-bottom or left-right search biases could not account for these trends.

Discussion.

The results of Experiment 4 show that perspective and compression gradients, both together and in isolation, elicit apparent depth effects early in visual processing. Depth effects found in the isolated conditions were smaller than in the combined gradient conditions, and at times emerged only in short target trials and displays with the fewest items. The impact on RT slopes is shown in Figure 15.

Insert Figure 15

The diminished apparent depth effect in the long target trials is easily accounted for by the reduced overall RTs. The reduced effect in the ten item trials is likely due to the item grouping on vertical trials, as discussed in Experiment 3. In this view, grouping may result from multiple items forming a plane separate from the background. While the items appear to float in their own vertical plane, it may be easy to ignore the background slanting away from the the search items and the viewer.

One additional influence from apparent depth emerged -- the type of background interacted with item orientation so that horizontal items most consistently produced the strongest apparent depth effects. The one exception was for short target - vertical item trials; when perspective was present on its own the apparent depth
Figure 15. Mean correct reaction time slopes (ms/item) in Experiment 4. Two search items (short and long) were examined against two locations (far and near in the slant condition, and left and right in the control condition) for each of six backgrounds (Perspective, Compression, Combined conditions and corresponding controls). Error bars are SEMs.
effects were reduced. This exception resembles Gillam's (1973) finding that size-scaling is largest when the orientation of the judged item is perpendicular to the gradient. The implication is that a large local contrast between the orientation of the item and surrounding gradient produces large apparent depth effects. Note that on the visual search conditions where this asymmetry in size-scaling emerged, the distractors were long, perhaps sufficiently so, to produce such a trend.

Regardless of the weakened effect in long target and multiple item conditions, and the additional interactions from item orientation, there remains plenty of evidence of apparent depth influences in the isolated gradient conditions. To assess the relative contribution of the dimensions I will separately consider (1) the magnitude of the effect, and (2) the consistency of the effects across different display sizes. The magnitude of the effect was greater for compression than perspective in the two item - short target condition. Compression showed a 50 msec mean difference across near and far locations for the two-item display while perspective showed only a 13 msec difference. However, the six-item trials in the compression conditions showed no location effect, and in the ten item trials there was a pronounced reversal in these trends, thus calling into question the reliability of the size-consistency effects in the compression conditions. If item grouping cannot account for the null or reversed location effects in the compression conditions, then the predicted direction of the effect would have to be argued to be most consistent on perspective trials (even with its smaller magnitude). A consistent apparent depth effect is also observed in long target trials with perspective, again demonstrating a slight advantage over compression in the supporting statistics.

Greater reliability in apparent depth trends on perspective trials may relate to its greater effectiveness in isolating distance information. Compression may be less reliable because of its tendency to confound orientation and distance information.
(Stevens, 1981), and to be confused with items that may differ in their size rather than their distance (Stevens, 1983b). To summarize, the evidence presented here shows that perspective may be more reliable across different conditions, but that the strength of the effect may be greater for compression. If grouping can account for the diminished effect in the compression conditions, then it is feasible that compression has both the more pronounced and the more reliable apparent depth effect.

Perhaps a better test of the relative influence of each dimension on search performance can be determined by comparing their influence when presented simultaneously. In this case, we are interested in how performance in the individual conditions predicts performance in the combined gradient conditions. Presumably, the dominant cue would be the better predictor of performance in the combined conditions. Such a test is presented at the end of Chapter 6.

Interestingly, Cutting and Millard (1984) and Vickers (1971) raise a point which may override any absolute hierarchical ordering of depth cue salience. The importance of the depth cues may be relative to the availability of alternative depth cues. Our visual system may simply use whichever information is available, and the contribution of each of the texture gradient dimensions to depth processing will change depending upon the information that is present at that time. This possibility is also considered in the next section and the final discussion where we consider how perspective and compression combine when both kinds of information are available.
Chapter 5

How does our visual system combine the dimensions?

The experiments presented up to this point have shown the influence of slant-from-texture on visual search performance. It has been proposed that mechanisms sensitive to projective geometry mediate this effect both early and late in visual processing. In addition, the individual texture gradient dimensions showed evidence for apparent depth effects on visual search performance. The present analysis extends this research by showing how these dimensions are combined by the early visual system. There are a number of approaches to understanding how perceivers combine multiple sources of information (e.g., Anderson, 1974; Ashby & Townsend, 1986; Garner, 1973; Sternberg, 1969a, 1969b). The framework used here combines a standard analysis of main effects and interactions derived from ANOVA (Anderson, 1974) with Sternberg's (1969a; 1969b) Additive Factor Method. Together, these allow us to assess whether early vision adds or multiplies perspective and compression and to infer whether there is a common code for the separate dimensions of the texture gradient.

Early vision can combine information from multiple sources in at least two general ways. One may involve specialized mechanisms early in the visual stream for the detection of particular surface properties. These detectors could be sensitive to information as simple as line orientation (Treisman, 1982; 1986) to as complex as shading gradients (Aks & Enns, 1992, Ramachandran, 1988), and may have evolved because of the ecological and functional reliability of these cues to scene properties (Gibson, 1966; Ramachandran, 1985). The idea of specialized detectors and corresponding trigger features can be traced to Barlow (1953) and can be seen in recent theories of object perception which postulate specialized mechanisms for volumetric solids (Biederman, 1987; Buffart, Leeuwenberg, & Restle, 1981; Leeuwenberg, 1988;
Pentland, 1986). The present results can be accommodated in a similar view by proposing that a *projective* mechanism is activated by the conjunction of appropriate texture gradient dimensions.

An alternative explanation is that early vision independently combines information from a number of sources. Cutting and Millard's (1984) finding of different factors responsible for the perception of flat and curved surfaces suggests that the perspective and compression gradients not only are independent sources of information, but they may be used for different purposes. Perspective may be used to detect the orientation of flat surfaces and compression may be used to detect the curvature in surfaces. The mechanisms that may detect this information could be similar to Stevens' (1981, 1983) depth and orientation map (see also Marr; 1982). Although Stevens (1981) differs from Cutting (1984) in his view of which computation is used to arrive at the estimates of depth and orientation (See Chapter 1 and Cutting, 1984), both agree that perspective provides the best information about distance and compression provides the best information about orientation or curvature.

One potential means for implementing multiple independent interpretations of unique image information could involve a large number of rapid and spatially-parallel processes that make "best guesses" about the scene based on information in the image (Enns & Rensink; 1990, 1991). These processes may be able to signal important scene properties collectively and stochastically. For example, one process might examine the perspective gradients in the image — a converging gradient would be interpreted as a surface receding away from the observer, whereas a diverging gradient would be interpreted as a surface slanting toward the observer. A second process might be involved in the analysis of compression gradients — a contracting gradient would be interpreted as a surface receding or curving away from the observer, whereas an expanding gradient would
be interpreted as a surface approaching the observer. Moreover, a first order change in compression may signal the presence of a flat slanted surface, while a second order change may signal the presence of a curved surface. Taken together, these processes would be able to signal the presence of important scene properties such as surface orientation (i.e., curvature) and distance (i.e., slant).

A test of additivity

In the test of additivity, the presence and absence of the texture gradient dimensions is varied orthogonally and subjected to an analysis of variance. This analysis can signal one of two broad strategies for combining information. First, an additive model would combine information by adding, subtracting, or averaging various sources. Second, a multiplicative model would combine information by multiplying or dividing. Additivity is operationalized in the ANOVA by the presence of main effects and the absence of interactions (Anderson, 1974). Interactions are most consistently associated with multiplicative models, although exceptions are possible (McClelland, 1979).

The present study also uses Sternberg's (1969a, 1969b) additive factors method (AFM) as a framework to further distinguish between an explanation of a specialized mechanism versus separate processes. A brief review of relevant assumptions of AFM are presented along with a description of the analysis and predictions (see also Taylor, 1976).

First, AFM assumes that successive and independent stages (or mechanisms) of processing intervene between the presentation of a search display and the subject's response. The relations between proposed stages of processing are established by selectively varying their temporal durations. In the present study, the duration of the potential apparent depth stage was manipulated by varying the type of texture gradient, the size of the target, and the apparent location of the target. The duration of an apparent depth stage is varied by the consistency of the target location and size on the apparently slanted
surfaces. Slowing of this stage is expected in the consistent conditions and speeding of this stage is expected in the inconsistent conditions. In addition, the usual visual-search variables of display size and target presence versus absence were varied to influence the number of items that needed to be inspected.

Second, if orthogonal variation in the difficulty of two factors such as perspective and compression leads to an additive pattern of performance, then the existence of two successive and independent stages of processing is implied. An additive RT pattern would be consistent with the operation of several processes where each make decisions on the basis of information available in the image. A detection decision in the visual search task would simply involve pooling the information from these independent processes. To the extent that decisions can be made quickly for each putative process, there should be a corresponding increase in search speed.

If, on the other hand, variation in the texture dimension results in an interactive pattern, then a common stage is implied. This would be most consistent with a specialized processor for slanted-texture surfaces. Combinations of texture gradient dimensions used by a specialized mechanism would influence search rate in a manner consistent with an interpretation of apparent depth (e.g., a perspective gradient in combination with a compression gradient would excite a slant-from-texture detector), whereas texture gradient combinations that were inconsistent would result in no apparent-depth influences on search (e.g., a vertical or horizontal line grid would not activate a slant-from-texture detector).

Finally, it is noted that AFM is not without its critics (e.g. Eriksen & Schultz, 1979; McClelland, 1979; Turvey, 1973), and that our experiment is not designed to compare AFM directly with alternative models. Nevertheless, these models all agree on the interpretation of an additive performance pattern: they all use this as diagnostic of separate stages. Where the models differ is in their interpretation of interactions. For
example, some models predict interactions from separate processing stages that overlap in time (McClelland, 1979). The path analysis presented in the final section helps distinguish between these potential interactive models.

**Method**

The first test of dimension additivity is based on a repeated measures ANOVA that includes perspective and compression as separate factors. The presence of each dimension was based on the data from the perspective, compression and combined conditions used in Experiment 4. The absence of perspective and compression were represented by the corresponding controls: horizontal and vertical grid conditions, and the absence of the combined gradients was represented by the average of the vertical and horizontal control conditions.

To aid in the presentation of these data, the main and interactive effects were expressed as contrast scores. The main effects from perspective (P), and compression (C), and the interaction from perspective and compression (PC) were represented by contrasts which combined the means of relevant dimension conditions. The means of the difference across dimension present and absent conditions were used to represent perspective: \([ (P1C1 - P0C1) + (P1C0 - P0C0) ] / 2 \), and compression \([ (P1C1 - P1C0) + (P0C1 - P0C0) ] / 2 \). The interaction between the dimensions was represented by the four contrasts that comprise the perspective and compression effects \([ (P1C1 + P0C0) - (P1C0 + P0C1) ] / 2 \).

An interaction contrast with a significant nonzero value indicates a reliable interaction between the texture dimensions. A positive interaction indicates that the combined dimension effects were larger than predicted by an additive model. A negative interaction indicates that the combined dimension effects were smaller than predicted by an additive model. Either type of nonadditivity would suggest that the dimensions influence
a common mechanism of processing (Sternberg, 1969).

Results

Figure 16 shows mean RTs and percent errors across the factors of perspective, compression, target size, apparent depth, and display size for target present trials.

Statistical analysis of these data included target size and apparent depth as a single factor -- size-depth consistency. Data were collapsed across trial blocks and missing values were ignored.

The presence of either or both texture dimensions made search more difficult when the target was present in the size-depth consistent condition (i.e., the short target was far or the long target was near). For each dimension of the texture gradient, performance was slower on consistent trials (size-depth consistency X perspective -- RT: $F(1,38) = 17.0, \ p < .001$, errors: $F(1,38) = .10$; and size-depth consistency X compression -- RT: $F(1,38) = 24.4, \ p < .001$, errors: $F(1,38) = 6.3, \ p < .05$).

These simple interactions along with the absence of their three-way interaction is the first suggestion of additive processing of the dimensions (RT: $F(1,38) = 2.4$, errors: $F(1,38) = .6$). Pairwise comparisons between the texture dimensions' present and absent consistent conditions were all significant ($t(38), \ p < .05$), except for the six and ten item short target trials; here performance was approximately the same for perspective present and absent trials. Additive trends were further supported by a significant main effect for perspective (RT: $F(1,38) = 8.9, \ p < .01$, errors: $F(1,38) = 8.4, \ p < .01$) and compression (RT: $F(1,38) = 10.9, \ p < .001$, errors: $F(1,38) = .8$), showing search was slower in the presence of either of the texture gradient conditions. Notice that the effect is proportionally the same for short and long target trials. The appearance of a smaller difference on long target
Figure 16a. Mean reaction time and percent errors across the factors of perspective (present, absent), compression (present, absent), and display size (2, 6, 10) for consistent size-depth conditions (short target - far, long target - near). Error bars are SEMs.
Figure 16b. Mean reaction time and percent errors across the factors of perspective (present, absent), compression (present, absent), and display size (2, 6, 10) for inconsistent size-depth conditions (short target - near, long target - far). Error bars are SEMs.
trials is a consequence of the overall lower RTs for these trials and the enlargement of the graph's scale to match the scale of the short target trials.

When the dimensions were presented together, subjects were much slower (i.e., overadditive on short target trials) in conditions with consistent size-depth information. These trends were supported by the interaction between perspective, compression, target size, and size-depth consistency (RT: $F(1, 38) = 7.8, p < .01$; errors ($F(1, 38) = 4.6, p < .05$). Inclusion of both texture dimensions in this interaction along with the noted idiosyncrasy of the six- and ten-item short target trials suggests that the visual system does not always engage in a purely additive strategy of processing texture gradient information. Instead, these trends indicate some dependence between processing the dimensions of the gradient in addition to independent processing occurring on long target trials. Perhaps these two strategies operate at different points in visual processing. This possibility is explored in further analyses of influences from target size, display size, item orientation, and processing time.

Contrast test of Additivity.

To further combine the effects of all conditions containing common texture dimensions, a contrast analysis was performed. The results are presented in Figure 17 where the approximately equal main effects of perspective and compression on search performance are immediately clear.

----- Insert Figure 17-----

Consistent with the earlier analysis, two combination rules emerged -- short target trials showed some interactive trends, and long target trials showed only additive trends. An analysis across the different display sizes confirmed that target size was not the sole predictor of additivity.
Figure 17. Mean difference in response time and percent error across all texture (perspective and compression) present and absent trials as a function of target size (short, long) and display size (two - white bars; six - gray bars; ten - black bars) for consistent size-depth conditions in Experiment 4.
**Processing speed and additivity.**

Assessing the relation between RT and additivity showed that processing speed may be a more reliable predictor of dimension additivity than either target size or display size. Recall that long target trials were substantially faster than short target trials as were trials with small display sizes relative to large display size trials (see Figure 16 & 17 and Tables 3, 5, & 6). Thus faster trials which included long target and small display size conditions show additive trends, whereas slower trials which included short target and large display size conditions show interactive trends. This relationship between processing speed and additivity was supported by a highly significant correlation, $r(5) = .84, p < .001$, between the PC interaction contrast and the average RTs that comprise the dimensions of all the contrasts. This correlation was based on mean values separated across target size and display size.

**Item orientation.**

Item orientation provided an additional influence on performance as shown in Figure 18 (perspective, compression, target size, display size X item orientation (RT: $F(2, 74) = 3.9, p < .05$). Search was slower on trials with horizontal items (RT: $F(1, 37) = 4.1, p < .05$, errors: $F(1, 37) = 9.3, p < .01$), especially in the consistent size-depth (item orientation X size-depth consistency -- RT: $F(1, 37) = 10.3, p < .01$, errors: $F(1, 37) = 9.4, p < .01$), and large display size conditions (item orientation X display size -- RT: $F(2, 74) = 7.3, p < .01$; errors: $F(2, 74) = 2.5$). Long target trials were most influenced by the orientation of the search items, with horizontal ones showing the greatest performance difference across consistent and inconsistent conditions (target size, size-depth consistency X item orientation -- RT: $F(1, 37) = 4.5, p < .05$, errors: $F(1, 37) = .05$).
Figure 18a. Mean reaction time and percent errors across the factors of perspective (present, absent), compression (present, absent), size-depth consistency (consistent, inconsistent) and display size (2, 6, 10) for vertical item conditions. Error bars are SEMs.
Size Depth Consistency

**Horizontal Items**

<table>
<thead>
<tr>
<th></th>
<th>DS2</th>
<th>DS6</th>
<th>DS10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short-Far</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT (msec)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Errors</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Long-Near**

<table>
<thead>
<tr>
<th></th>
<th>DS2</th>
<th>DS6</th>
<th>DS10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RT (msec)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Errors</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Perspective**

Figure 18b. Mean reaction time and percent errors across the factors of perspective (present, absent), compression (present, absent), size-depth consistency (consistent, inconsistent) and display size (2, 6, 10) for horizontal item conditions. Error bars are SEMs.
The contrast analysis confirmed these effects from item orientation. In addition, the contrast analysis shown, in Figure 19, suggests approximately equal contributions from the texture dimensions in both vertical and horizontal trials. There are only two exceptions. First, the contribution from perspective is slightly weaker when items are oriented vertically. This however could be the result of a floor effect in long target trials, with the effect of reducing apparent depth influences. Second, compression contributes more than perspective in short target - horizontal trials when two items were in the display.

More pertinent to the question of additivity, the evidence is overwhelming for the absence of interactions in long target trials and the presence of interactions on large display size - short target trials, regardless of item orientation. Even though size-depth consistent effects were larger on horizontal trials, the pattern of additivity was unaffected.

**Speed of Processing, additivity, and item orientation.**

The relationship between processing speed and additivity was again supported by strong positive correlations between the compression by perspective interaction contrast and mean RTs of the conditions making up the contrasts (vertical trials: $r(5) = .90$, $p < .001$; horizontal trials: $r(5) = .64$). Low statistical power is likely causing the absence of significance in the horizontal condition correlation.

**Discussion**

The tests of additivity, together with correlations between average RT and combined dimension RT, provided evidence for independent coding of two dominant dimensions of a texture gradient in fast visual search conditions and shared coding in slow search conditions. These findings support the view that separate visual pathways operate on the two dimensions at the earliest stages of visual processing and that their output may feed into a single pathway later in visual processing.
Figure 19a. Mean difference in response time and percent error for short target conditions across all texture (perspective and compression) present and absent trials as a function of display size (two - white bars; six - gray bars; ten - black bars) and item orientation (vertical and horizontal).
Long Target

Vertical Items

Horizontal Items

Display Size

\[ \begin{array}{c}
\square & 2 \\
\black & 6 \\
\black & 10 \\
\end{array} \]

Background

\[ \begin{array}{c}
P \\
C \\
PC \\
\end{array} \]

\[ \begin{array}{c}
P \\
C \\
PC \\
\end{array} \]

\[ P = \text{Perspective} \]
\[ C = \text{Compression} \]
\[ P \times C = \text{PC interaction} \]

*Figure 19b.* Mean difference in response time and percent error for long target conditions across all texture (perspective and compression) present and absent trials as a function of display size (two - white bars; six - gray bars; ten - black bars) and item orientation (vertical and horizontal).
Evidence for additive processing is consistent with a number of findings and theories in the texture gradient literature. First, the present findings compliment those of Cutting and Millard's (1984) who suggest that different performance for flat and curved surfaces implies that the perspective and compression gradient are independent sources of information and are coded independently in vision. Vickers (1971) also has shown that slant and depth impressions become additively stronger with increases in the amount of available perspective and compression information. Possible computational instantiations of such independent coding include Marr's (1982) and Stevens' (1981; 1983) depth and orientation maps, to which we will return to later in this discussion.

Interactions between the dimensions found in later visual processing require an alternative mechanism that integrates the information from these earlier distinctive codes. One candidate, suggested earlier, may be mechanisms receptive to fairly complex information used for the detection of particular surface or object properties. Sensitivity to a shading gradient, for example, would involve detecting the direction of a luminance gradient, perhaps reflecting early sensitivity to 3-D curvature and the direction of light in the scene (Ramachandran, 1988). However, further research including the present results show some limits to the specificity and complexity of the initial coding of shading (Aks & Enns, 1992) and texture gradient information. At the earliest stages, the individual dimensions, on their own, may activate a mechanism sensitive to depth information and at least for the texture gradient, that scales for size. Sensitivity to the combination of the dimensions occurs slightly later in visual processing with the texture gradients triggering an even stronger signal in the size-scaling mechanism.

There are at least two interpretations for the additive effects of perspective and compression at the earliest stages of visual processing and the interactive effects occurring later in visual processing. One possibility is that the dimensions are registered
in parallel at two separate sites, each with its own function -- one for orientation and one for distance (i.e., Marr, 1982; Stevens, 1983). Although the output of one does not depend on the other there may be some overlap in the information that each map codes -- such redundancy is advantageous in an environment where information is known to vary. After the separate computations are completed they may combine later in a mechanism that provides a full representation of surface shape. For this interpretation to be valid, we would expect that perspective and compression should independently predict performance in the combined conditions and that their temporal coding order is not contingent upon each other. This possibility is explored in a final analysis of the processing sequence.

An alternative explanation of the additive findings is that only one dimension may be processed at a time, since the output of one mechanism might improve the output of the other. Here we might expect a dependency in the processing of dimensions which may in turn be reflected in sequential processing of perspective and compression. Such a view allows for prioritizing the use of these mechanisms for particular (but redundantly coded) tasks. For example, since perspective codes only distance information and compression combines orientation and distance in its code, it would make sense for the visual system to first use perspective to establish the distance to a surface rather than rely on potentially erroneous data from the confounded compression information. Compression may be used later to add information about surface orientation and shape.

Both parallel processing or perspective-first models are consistent with the theories of Stevens (1981; 1983) and Marr (1982), and provide concrete predictions about processing sequence. The discernible nature of perspective and the ease of extracting distance information from it (see discussion on the characteristic dimension), makes it a likely candidate for being processed prior to compression. Accordingly,
Slant - from - texture

perspective should better predict performance in the combined condition regardless of whether compression is available. The plausibility of these alternatives are tested in an assessment of the temporal sequence of dimension processing.

Processing sequence.

To assess the temporal order of dimension processing a path analysis derived from correlations between the texture conditions was performed. The technique is based on Wright (1921), Simon (1953), and Blalock (1962; 1985); a brief summary and application to illusion magnitude is presented in Coren and Ward (1979). This path analysis allows us to distinguish between an interpretation of parallel and sequential processing of the texture gradient dimensions and to assess the independence of processing the separate texture dimensions. Earlier results from the combined conditions show that spatial processing occurred in parallel -- RT slopes were relatively shallow in both size-depth consistent (7.1 msec/item) and inconsistent trials (8.7 msec/item). However, spatial processing of search items need not relate to either the spatial or temporal processing of the underlying texture. Locating an item among related items may be a distinct process from one that registers background depth information. Similarly, additivity found in Experiment 4 does not inform us as to whether the underlying processes operate temporally in serial or parallel. Hence the path analysis.

Method

A path analysis based on correlations between RTs from the separate dimension [i.e., perspective (P) and compression (C)], and combined conditions [i.e., perspective and compression (PC)] was performed to determine the order of dimension processing. According to this analysis, a nonsignificant correlation between P and C (rpc), and significant correlations between the individual and combined dimensions [rp(pc); rc(pc)]
indicate a parallel model. A finding of all significant correlations \([r_{p,c}; r_{p(pc)}; r_{c(pc)}]\) indicate a sequential model.

Partial correlations were used to further determine the order of dimension processing and to indicate the interdependence of the individual dimensions in predicting the combined performance. A nonsignificant partial correlation: \(r_{p(pc)}c\), where compression is controlled, with a significant correlation: \(r_{c(pc)}p\), where perspective is controlled, indicates a dependence of perspective on compression and imply a processing order of P -> C -> PC. In other words, removing the influence from compression eliminates the effectiveness of perspective predicting search performance in the combined gradient condition. A nonsignificant partial correlation: \(r_{c(pc)}p\) with a significant, \(r_{p(pc)}c\) indicate a dependence of compression on perspective and implies a processing order of C -> P -> PC. Notice that an asymmetry in the significance of the partial correlations indicates the path's direction.

**Results**

Correlations between RTs from the separate and combined dimension conditions are presented in Table 7. In the long target trials, perspective was the only significant predictor of combined gradient performance \((r_{p(pc)} = .43, p < .01)\). The non-significant correlation between perspective and compression \((r_{pc} = .22, p = .17)\) indicates that the individual dimensions, if indeed both are coded, are coded in parallel. However, the absence of a significant correlation between compression and the combined gradients, \((r_{c(puc)} = .08)\) indicates that the parallel model reduces to one where perspective alone predicts performance in the combined trials, as shown in Figure 20.

\[\text{Insert Table 7 and Figure 20}\]

In the short target trials, correlations between perspective and compression were significant \((r_{pc} = .40, p < .01)\), as were each of their correlations with the combined
Table 7

Intercorrelations and partial correlations among background conditions: Perspective (P), Compression (C), and Perspective & Compression (PC)

<table>
<thead>
<tr>
<th></th>
<th>Short</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC</td>
<td>P</td>
</tr>
<tr>
<td>Correlations</td>
<td>PC</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>.37*</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>.45**</td>
</tr>
<tr>
<td>Partial</td>
<td>PC</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>.35*</td>
</tr>
</tbody>
</table>

* p < .05
** p < .01
**Performance Speed**

**Fast**

P → PC

**Slow**

P → C → PC

---

**Figure 20.** Results of path analysis of texture gradient dimensions separated by performance speed. Fast performance in the combined gradient conditions is predicted best by perspective alone. Slower performance is represented best by a sequential process with perspective coded prior to compression. Partial correlations between the texture gradient conditions are reported.
gradient trials ($r_{tp}(pc) = .37, p < .05; r_{tc}(pc) = .45, p < .01$). The three significant correlations indicate that a sequential model is likely to be operating. An assessment of the partial correlations, $r_{p(puc)}c = .22, n.s.$, and $r_{c(puc)}p = .35, p < .05$, indicates that perspective reliably predicts combined performance only when compression information is also present. Given the asymmetries in the partial correlations together with the significant correlations between the isolated dimensions and combined conditions for the short target trials, the implied order of processing requires perspective to precede compression as shown in Figure 20.

In an analysis of the relation between RTs and processing sequence, short target trials -- also the slower conditions (RT > 650 msec) -- showed evidence for sequential processing. One might expect this trend to be reflected in processing of the separate dimensions (i.e., one dimension should be processed prior to another even when presented separately). However separate dimension conditions did not show this. Instead, the texture gradient dimensions were processed in an equal amount of time in trials where they were presented on their own. Mean RT for perspective was not significantly faster than compression on short target (perspective = 679 msec, compression = 682 msec) and long target trials (perspective = 600 msec, compression = 606 msec), reflecting no difference in RTs for the interactive - sequential strategy, as well as the additive-parallel processing strategy. Keep in mind that these RTs were extracted from the separate texture conditions and that processing strategies can (and appear) to change when new information is present. Also note that asymmetrical correlations do not necessarily reflect a temporal sequence but instead may indicate intervening mechanisms (Blalock, 1985). In this case the dependency of perspective on compression can be interpreted as compression simply intervening in the processing sequence, perhaps providing additional information about surface orientation or shape.
Discussion

In the present study, only perspective information was needed in fast search conditions to predict performance in combined texture conditions. In slow search trials, both texture dimensions were needed, with the restriction that perspective was coded first. To account for these performance trends let us return to the two models described earlier -- the parallel processing and the perspective coded first model. Parallel processing is supported by uncorrelated influences between the two texture dimensions on search performance, and implies independent coding at two separate sites, each with its own function -- one for orientation and one for distance. According to this view, the output of one does not depend on the other but there still may be some overlap in the information that each map codes (i.e., perspective and compression code distance information). The output of the separate computations may combine later in a mechanism that provides a more complete description of surface shape.

Although perspective was largely unaffected by the presence of compression in the fast search conditions, the overall pattern of results from the path analyses did not support the parallel processing view. Instead, trends were most consistent with the alternative -- perspective first model. According to this view and the results of the path analysis, only one dimension is processed at a time, and the output of the perspective mechanism contributes interactively to the output of the compression mechanism. The fast trials reflect influences only from the perspective segment of the processing sequence. The slower trials, which are also driven by early vision, last long enough to capture the entire sequence of perspective then compression processing.

A number of interesting questions arise regarding the characteristics of processing in fast and slow conditions and what function they might serve; two are considered here.
Why does early vision need both components on slow trials while one is sufficient on fast trials? and Does sequential processing imply a priority of processing?

One of the more plausible answers to the first question is that slow trials may require additional information because of the difficulty involved in detecting the short target, especially in the presence of many distractors. The difficulty of short target trials relative to long target trials is likely due to the search asymmetry that occurs reliably in any task involving size-discrimination -- finding a small object among large ones (of the same type) is almost always more difficult than finding a large object among small ones. Here it is argued that detection of a small object (or an object located among many items) may be improved by adding information about its orientation. Results of the path analysis, together with those of Experiments 3 and 4, showed that the texture gradient elicited a stronger apparent depth effect when compression was used with perspective. With perspective yielding information about distance, the extra orientation information from compression is likely to improve estimations of object distance and size in the 3-D environment.

An answer to the second question about the sequence and priority of processing follows from an explanation of why more information may be used in slow search conditions. Granted, the act of combining the texture dimensions may constrain our visual system to register information sequentially. But the question of priority remains. Logically, the donor process must come first -- that is the process whose output supplements a subsequent process. Perspective would be expected to have processing priority for a number of reasons -- its greater ecological validity (Coren, personal communication), its greater reliability in signaling distance information (Stevens, 1984), or it may simply be easier to extract from the image (Stevens, 1981). Thus it seems plausible that early vision extracts distance first from perspective, and afterwards, this
information is used to isolate orientation from the compression gradient. Certainly, knowing the correct distance and orientation of an object helps us decipher ambiguous information and make better judgments about other object properties (e.g., size).

**General Discussion**

The existing literature on the perception of depth from texture is inconclusive regarding how subjective depth impressions relate to rapid and spontaneous depth detection. The present set of experiments shows that there are strong parallels between the texture information used in phenomenal perception and speeded performance. Differences that do occur appear to be in the way in which the information is combined. The present studies show independent coding of the perspective and compression dimensions at the earliest stages of visual processing and integrated coding later in visual processing. Presumably this pattern reflects a progression of visual processing from initial coding in distinct neural subsystems to a synthesis in a common system that leads to our final unitary percepts. A summary of the results leading up to this conclusion is presented, followed by a discussion of the implications for theories of early vision and scene perception.

The first two experiments confirmed that subjective perception of slant-from-texture predicts speeded performance in a visual search task and that the mediating representation relies on principles of projective geometry. The third experiment showed further that even *early* vision is sensitive to this apparent depth information in the texture gradient. However, one unexpected finding was that trials with many items showed the weakest apparent depth effects. This result was attributed to a novel grouping phenomenon induced by a large number of same-size items against the texture gradient, allowing subjects to ignore the apparent depth information in the background. The fourth experiment showed that both dimensions of the texture gradient are detected by
early vision. Stronger apparent depth influences were found in trials where both perspective and compression dimensions of the texture gradient were available, and when search items were oriented horizontally so that they appeared to be attached to the underlying gradient. The fifth experiment assessed how early vision combines the dimensions of the texture gradient. The texture dimensions were shown to have independent influences early in visual processing. However, subjects showed greater sensitivity to the interaction of the dimensions during slower search conditions. These findings support the view that perspective and compression are coded separately at the earliest stages of visual processing and are brought together only later.

Implications for theories of early vision.

Recall from the discussion in Experiment 3 that researchers who use the visual search test typically regard early vision as an analytical process where a limited set of features are coded in parallel (Beck, 1966, 1967; Julesz & Bergen, 1983; Treisman, & Gelade, 1980; see also theories of surface perception below). The data presented here question this strict view. Early vision's sensitivity to texture information, along with its corresponding apparent depth interpretation, contradicts this view. Each dimension, even on its own, conveys visual information that when viewed from a conventional perspective would be defined as a conjunction of simple geometric features -- the radial perspective pattern consists of multiple orientations, and the foreshortened compression arrangement contains a range of different sizes. These experiments suggest that early vision may detect an alternative form of information such as the constant variations (derivatives) in size and orientation in the texture gradient. However, these have not been classified as primitive features (Northdurft, 1985; Treisman & Gelade, 1986), nor are they easy to estimate (see Stevens, 1981, 1984).
Recent studies, however, show support for greater processing sophistication at the earliest stages of visual processing. These include studies that show sensitivity to binocular disparity and motion (McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986), spatial line relations and shading, (Enns & Rensink, 1990a, 1990b), and direction of lighting source and luminance gradients (Aks & Enns, 1992; Kleffner & Ramachandran, 1992; Ramachandran, 1988). In all of these studies early vision recovers 3-D scene properties which, when described in terms of the image, are complex conjunctions. The present study also shows early encoding of a conjunction (or 3-D representation) and, in addition, size-scaling at the earliest stages of visual processing.

It is important to note that different performance in the near and far locations of the present study may have been a consequence of the texture gradient or the local contrast of the items' size relative to the size of the surrounding background. The latter influence would reflect a strategy where subjects compare the size of the target with its local background. Detecting a large item against small texture elements may be easier than detecting a large item against large texture elements. One way to distinguish between a search strategy based on detection of the local contrast versus the gradient would be to compare search trials with small texture elements in the background relative to trials with large elements in the background. Smaller performance differences across these conditions relative to the differences across the locations of the gradient conditions would suggest that the global gradient, and not just local contrast, is responsible for shifts in performance.

Even if local contrast rather than global gradient information is responsible for performance trends, additional apparent depth effects did emerge. These included top-bottom influences resembling a height-in-plane interpretation, large apparent depth
effects when items were oriented horizontally so that they were consistent with the slanted surface, and item grouping in the ten-item condition with its subsequent item-background segregation. In all of these cases, apparent depth appears to have influenced search performance. The same may also be true of a search strategy based on local contrast. Early assumptions of projected size — small background texture elements signaling greater distance than large elements — may be applied to this information either when it is presented as part of a gradient or as a homogeneous background.

Regardless of whether local contrast or gradient information governs these trends, early and late systems both appear to have been influenced by the same information. This sensitivity to common information poses further problems for distinguishing stages of texture processing. At least with respect to texture gradient information, the difference between the two stages may not be in what information is used, but in how it is combined. Perhaps a more discerning diagnostic for early and late visual processing of texture gradient information may be based on how the systems code multiple sources of information. The trends presented here suggest separate encoding for early and integrated encoding for late processing.

**Implications for theories of surface perception.**

Similar to the analytic view of early visual processing, independent coding for various depth cues has also been proposed (Attneave, 1972; Marr & Nishihara, 1978), and supported by psychophysical demonstrations (Bruno & Cutting; 1988; Cutting & Millard, 1984; Stevens, Lees, & Brookes, 1991). When the texture dimensions are classified as depth cues, rather than as a collection of simple-geometric features, evidence for independent processing of the texture dimensions agrees with the analytic views of early vision researchers and findings in the field of depth perception. What
has been neglected in the field of depth perception is a description of how early the independent processing occurs.

How compatible are the present results with models of depth perception? The appropriate models to assess are those that concentrate on surface perception. Most well known is Marr's (1978, 1982) three stage model, which includes an important distinction between distance and surface orientation at the intermediate stage of the 2 1/2-D sketch. The 2 1/2-D sketch describes surfaces visible from a given viewpoint, with values of distance and surface orientation metaphorically represented by thousands of "needle-like" points. Orientation is coded by the direction of each surface normal (i.e., the needle) as projected on this surface, and slant is indicated by the length of each needle (e.g., zero length indicates no slant). Marr's model assumes a gradient space corresponding to each point in the visual field, which can provide a full description of surface shape in a local region.

Stevens (1981; 1983) has elaborated Marr's model by describing the most efficient information and computations necessary to give rise to such a surface mapping. Briefly stated, distance information is most easily extracted from the perspective dimension (or equivalently the characteristic dimension) of the texture gradient. Perspective is simply the direction of least texture variation, and distance is the reciprocal of this scale. To extract slant, there are a number of potential solutions including coding the gradient of the CDs, detecting various surface derivatives, or performing any one of a number of trigonometric analyses. Although a number of these solutions are plausible, none has been shown to be the one used in human vision. Also, many are prone to error as they tend to confound information about distance and orientation. Note that for the size-discrimination task used here, only distance is needed. When perspective is available, distance can be directly extracted from this information. When compression information is presented on its own, it
is easier to extract distance after calculating surface slant or orientation. In accounting for
the apparent depth effects of the present study, the most parsimonious account is likely to
involve scaling for size to adjust for early assumptions of projected size and perceived
depth. The present results, together with Marr’s and Steven’s theoretical accounts, show
that particular dimensions may be more prone to size-scaling (i.e., perspective) but when
information is limited we use what is available (i.e., compression).

**Early vision and surface perception.**

A complete model of early vision must be able to explain how we perceive surfaces
specified by texture gradients. Such a model must unify data from two traditionally
distinct fields of perception -- early vision and surface perception. Each of these fields
must be able to account for various aspects of the present findings.

*First*, theories of early vision (i.e., Treisman, 1986) and surface perception (Marr,
1982) already both propose that vision must have early independent codes that can be
activated rapidly and in parallel. *Second*, at least some of these early codes must respond
to combinations of features that signal important information about properties of surfaces
in the environment. Only recently has this been argued to be an important aspect of early
vision (Enns & Rensink, 1990a, 1990b; Ramachandran, 1988). The present study
contributes to this suggestion in showing that a radial and foreshortened pattern of lines
can also signal apparent depth information early in vision. *Third*, these codes must be
sensitive to viewer-centered information and have intrinsic scaling properties. The first
part of this criterion is consistent with Epstein’s theory that preattentive vision is sensitive
to projective information (Epstein & Babler; 1989, 1990). However, the addition of size-
scaling mechanisms, at this early stage, has yet to be incorporated into theories of early
vision. Theorists must heed the clear evidence that even simple size judgments, driven by
early vision, compensate for surrounding depth information. *Finally*, these codes are
integrated over time (as exposure duration is lengthened) to enhance depth signals and assumptions of projective geometry.

Existing models of early vision and scene perception (i.e., Enns & Rensink, 1990, 1991; Epstein & Babler, 1989; Grossberg & Mingolla, 1985; Marr, 1982; Treisman, 1986) need to include these essential components to provide a more complete account of human vision. Perhaps most important is the need to incorporate early coding of separate texture dimensions so that depth information is intrinsic to each code. In addition, since both visual systems are sensitive to the same texture information, theorists should consider that early and late coding of surface properties may be best distinguished on the basis of how fundamental dimensions are coded. When time is limited, early vision codes many sources of information independently, perhaps even selecting the most direct signals for surface properties. Only later in vision, when more time is available, does a unified code become available to integrate these separate properties of the three-dimensional world.
References


Appendix A

Experiment 1

Target presence versus target absence.

Visual search on target present trials was substantially faster and less accurate than target absent trials (RT: $F(1,10) = 36.6, p <.001$; and errors: $F(1,10) = 8.5, p <.05$). RTs increased with display size (RT: $F(2,20) = 69.7, p <.001$; errors: $F(2,20) = .4, n.s.$) primarily for target present trials as confirmed by significant interaction of display size with target presence (RT: $F(2,20) = 34.6, p < .001$; errors: $F(2,20) = 3.3, n.s.$).

Experiment 2

Target presence versus target absence.

Visual search on target present trials was substantially faster and less accurate than target absent trials (RT: $F(1,18) = 49.5, p <.001$; and errors: $F(1,17) = 26.7, p <.05$). RTs increased with display size (RT: $F(2,36) = 114.4, p <.001$; errors: $F(2,34) = 1.4, n.s.$) primarily for target present trials as confirmed by the significant interaction of display size with target presence (RT: $F(2,20) = 34.6, p < .001$; errors: $F(2,20) = 3.3, p <.01$).

To ensure apparent depth effects were due to the presence of a target in apparent locations rather than to configural effects of the display, target absent trials were analyzed by location. This was possible since the arrangement of target absent trials was the same as target present trials with the exception that the target was replaced by a distractor item on the absent trials. These distractors, therefore, behaved as an additional control for target present trials. Evidence for apparent location trends occurring only in target present trials were supported by non-significant differences across all locations of the target absent trials (All Fisher's LSD tests, $p > .05$). Further support for slant-from-texture effects being unique to target present trials include the significant interaction between slant, target size, target location and target presence (RT: $F(2,36) = 28.2, p <.001$; and errors: $F(2,34) =$
22.9, p < .001), and the three-way interaction between target location, target size, and target presence for RT, F(2,36) = 41.3, p < .001; and errors, F(2,34) = 29.7, p < .001. The restriction of slant effects to target present trials was confirmed by the interaction between target presence and slant for errors, F(1,17) = 25.4, p < .001; and the interaction between target presence and location for RT, F(2,36) = 51.8, p < .001; and % errors, F(2,34) = 28.1, p < .001.

**Experiment 3**

**Target presence versus target absence.**

Visual search on target present trials was substantially faster and less accurate than target absent trials (RT: F(1,39) = 104.9, p < .001; errors: F(1,40) = 135, p < .001). RTs and errors increased with display size (RT: F(2,80) = 174.8, p < .001; errors: F(2,80) = 12.0, p < .001), resulting in mean RT slopes of 7 msec/item for target present trials and 13 msec/item for target absent trials. This RT Slope difference was supported by a significant interaction between display size with target presence (RT: F(2,80) = 31.3, p < .001; errors: F(2,80) = 7.99, p < .001).

**Practice**

A significant main effect of trial block was observed (RT: F(2,78) = 14.8, p < .001; Error: F(2, 80) = 2.7, n.s.) which interacted with target presence (RT: F(2,78) = 3.4 p < .05; error: F(2, 80) = .29, n.s.). This block X target presence interaction is a result of RT performance showing the greatest improvement over blocks on target absent trials.

**Experiment 4**

**Perspective.**

Similar to the combined gradient conditions, performance on target present trials was substantially faster and less accurate than target absent trials (RT: F(1,40) = 106.9, p < .001; errors: F(1,38) = 135.3, p < .001). RTs (F(2,80) = 117.4, p < .001) and
percent errors ($F(2, 76) = 12.8, p<.001$) increased with display size resulting in RT slopes of 6.6 msec/item for target present trials and 8.3 msec/item for target absent trials. This RT Slope difference was supported by a significant interaction between target presence and display size for RTs, ($F(2, 80) = 8.3, p<.001$), and errors, ($E(2, 80) = 34.2, p<.001$).

**Practice.**

A significant main effect of trial block was observed for RT ($F(2, 80) = 7.4, p<.001$), but not for errors ($F(2, 76) = 0.5, p<.001$). Most of the improvement in RTs is accounted for by target absent trials ($F(2, 80) = 5.6, p<.01$). Fewer errors were made on intermediate display size trials in the first block’s slant condition, and in the third block’s no slant condition. These trends were supported by a significant interaction between trial block, background, and display size ($E(4, 152) = 2.8, p<.05$).

**Compression.**

Performance on target present trials was substantially faster and less accurate than target absent trials (RT: $F(1,39) = 137.4, p<.001$; errors: $F(1,38) = 79.9, p<.001$). RTs ($F(2,78) = 121.6, p<.001$) and percent errors ($F(2, 78) = 121.5, p<.001$) increased with display size resulting in RT slopes of 6.4 msec/item for target present trials and 11.7 msec/item for target absent trials. This RT Slope difference was supported by a significant interaction between target presence and display size (RT: ($F(2, 78) = 20.0, p<.001$; errors: ($F(2, 76) = 16.3, p<.001$).

**Practice.**

Subjects showed a general improvement in performance in the compression condition as reflected in the significant trial block effect for RT ($F(2, 78) = 22.5, p<.001$), but not for errors ($F(2, 76) = 0.5, n.s.$). Most of the improvement occurred in trials where the target was on the left or right of the six or ten item displays (trial block, location X display size ($E(4, 152) = 2.5, p<.05$).