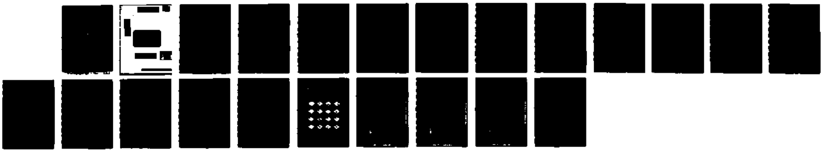
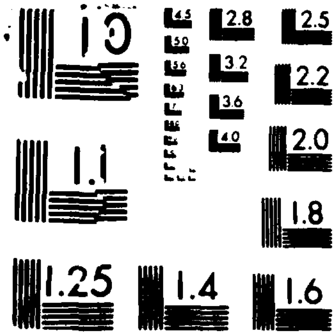
AD-A181 829 THE KINETIC DEPTH EFFECT AND IDENTIFICATION OF SHAPE 1/1
(U) NEW YORK UNIV NY COURANT INST OF MATHEMATICAL
SCIENCES M S LANDY ET AL JAN 87 TR-87-1
UNCLASSIFIED N00014-85-K-0077 F/G 12/2 NL




1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

Department of Computer Science
urant Institute of Mathematical Sciences
251 MERCER STREET, NEW YORK, N.Y. 10012

11

The Kinetic Depth Effect and Identification of Shape

by

Michael S. Landy †
George Sperling †
Barbara Dosher ‡
Mark E. Perkins †
†Psychology Department, NYU
‡Psychology Department, Columbia University

Mathematical Studies in Perception and Cognition
Technical Report 87-1
January, 1987

†Psychology Department
New York Univesity
6 Washington Place
New York, NY 10003 USA

The work described in this paper was supported in part by a grant from the Office of Naval Research, Grant N00014-85-K-0077, Work Unit NR 4007006.

Mathematical Studies in Perception and Cognition

87 - 1

The Kinetic Depth Effect and Identification of Shape

Michael S. Landy[†], *George Sperling*[†], *Barbara A. Dosher*[*], *and Mark E. Perkins*[†]

[†]Psychology Department
New York University

and

[*]Psychology Department
Columbia University

Running Head: KDE and Shape

Keywords: KDE, Kinetic Depth Effect, Structure From Motion, Shape

Address correspondence to:

Michael S. Landy
Psychology Department
New York University
6 Washington Place, Rm. 961
New York, New York 10003

ABSTRACT

This paper introduces a new method for assessing the effectiveness of kinetic depth stimuli for creating a percept of three-dimensional shape. The task is shape and motion identification, where each shape presented is one of a large lexicon of shapes. The shapes consist of bumps and depressions on an otherwise flat ground. They vary in the number of bumps, their position and size. Using multi-dot representations of the shapes, identification is demonstrated to increase with dot numerosity and with the extent of depth portrayed. This task holds promise as a paradigm for examining objectively the cues necessary for the kinetic depth effect. Accurate performance on the task requires a global percept of three-dimensional shape, and is not prone to subject strategies using simple velocity measurement at a small number of spatial locations.

INTRODUCTION

In 1953, Wallach and O'Connell introduced the notion of a depth percept derived purely from relative motion cues, which they called the 'Kinetic Depth Effect'. Since that time, there has been a great deal of research on the problem, examining the effects of stimulus parameters such as dot numerosity in multi-dot displays (Green, 1961; Braunstein, 1962), frame timing (Petersik, 1980), occlusion (Andersen & Braunstein, 1983; Proffitt, Bertenthal, & Roberts, 1984), the detection of non-rigidity in the three-dimensional form most consistent with the stimulus (Todd, 1982), veridicality of the percept (Todd, 1984a,b), etc.

At the same time, there have been several attempts at modeling how observers derive three-dimensional structure from two-dimensional motion cues. Ullman (1979) referred to this computational task as the 'Structure from Motion' problem. Several models are essentially geometry theorems concerning the minimal number of points and views needed to specify the shape under various simplifying assumptions such as rigidity (Ullman, 1979; Webb & Aggarwal, 1981; Hoffman & Flinchbaugh, 1982; Hoffman & Bennett, 1985; Bennett & Hoffman, 1985). A few models make use of measurements of point velocity (i.e. an optic flow field) in addition to point position (e.g., Clocksin, 1980; Longuet-Higgins & Prazdny, 1980; Koenderink & van Doorn, 1986), and one also uses point acceleration (Hoffman, 1982). Finally, there are process models which utilize changing relative position data as they develop a three-dimensional representation, while attempting to minimize departures from rigidity in that representation (Ullman, 1984; Landy, 1987).

It has been difficult to relate models of the KDE to the results of psychological studies. Part of the problem has been the difficulty of finding an appropriate experimental paradigm. Many KDE experiments have used subjective ratings of 'depth' or 'rigidity' as the response. Relating such a subjective response to a process model is problematic.

Another approach is to test the accuracy of the KDE in an objective fashion. Does the observer perceive the correct depth? The correct depth sign? The correct depth order? The correct curvature? The studies cited above have attempted to answer many of these questions using objective response criteria (e.g., percent correct in a one- or two-interval forced-choice task). Unfortunately, in almost every case, reasonable subject performance on the task is possible without the subject actually perceiving depth. This is because, in each case, there exists a simple local cue sufficient to make the judgement accurately.

Let us examine some examples. In the study by Lappin, Doner, & Kottas (1980), subjects are required to determine which of two two-frame displays has a higher signal-to-noise ratio (in terms of dot correspondences). The signal dots represent two frames of a rigid rotating sphere, but this fact is not relevant to the response, which only requires determining the percentage of dot correspondences consistent with a particular optic flow field (that of the rotating sphere).

In two studies by Petersik (1979, 1980), the task is discrimination of rotation direction, where polar perspective is used. The difficulty here is that the motion of a single dot is sufficient to respond correctly. Under polar perspective, stimulus points follow elliptical paths in the image plane. To determine rotation direction, the subject need only determine 2D rotation direction of a single point (assuming knowledge of the vertical position of the point with respect to eye level). Braunstein (1977) examined this point specifically, and determined that only the vertical component of the polar perspective transformation was used by subjects for such a judgement.

Andersen & Braunstein (1983) also use rotation direction discrimination. In their displays, parallel perspective is used. The cue to depth order is provided by occlusion (regions in the front surface occlude points on the back surface). Again, subjects do not need to perceive a 3D object to perform the task. A subject need only determine whether, say, leftward moving points are continuously visible or not.

In several studies, simple relative velocity cues are all that is required to perform the given task. In Braunstein & Andersen (1981), a multi-dot display of a translating dihedral edge is presented. Subjects judged whether a given display represented a convex or concave edge. In this task, comparing the relative velocity of points in the center and at the top edge of the display at a fixed point in time is all that is necessary to perform accurately in the task. In experiments by Todd, subjects determine which of five curvatures (Todd, 1984a) or slants (Todd, 1984b) are depicted in a multi-dot display. The task is again described in terms of the 3D object perceived, but accurate performance is possible by comparing the relative velocities of points in two areas of the display. Finally, in experiments by Inada, et al (1986), subjects view displays of three points rotating in depth and are to determine which point has the intermediate depth value when the display terminates. This task is again subject to simple velocity computations not requiring knowledge of the depth portrayed. For example, if the axis of rotation is in the image plane, the point with the intermediate depth will nearly always be the point with the intermediate 2D velocity.

One possible solution to these problems is to prevent subjects from using anything but the perceived 3D shape by not providing feedback. This approach has been used extensively by Todd (1982, 1984a, 1984b). Unfortunately, withholding feedback brings along its own problems, such as subject bias.

The problem is this: The KDE is a perceptual phenomenon which allows subjects to perceive the relative depth of different positions in visual space, and hence to infer the shapes of objects in the environment. None of the experiments discussed above require the subject to have perceived a 3D shape in order to perform accurately.

In this paper, we describe a new method for investigating the kinetic depth effect. The task is shape identification, where on each trial, one of a large lexicon of shapes is presented. Each shape consists of a flat ground with zero, one, or two bumps or depressions. The bumps and depressions vary in position, and in two-dimensional extent. Because of the way in which the lexicon is constructed, a global perception of shape is required for good performance. Simple subject strategies involving a small number of local measurements do not suffice to carry out the task. We report here a use of this new experimental paradigm to investigate the effects of dot numerosity and depth extent on the effectiveness of the KDE.

METHOD

Subjects. Three subjects were used in the study. Two are authors, and the third was a graduate student naive to the purposes of the experiment. All had normal or corrected-to-normal vision.

Displays. The shapes used in the experiment were three-dimensional surfaces consisting of zero, one, or two bumps or concavities on an otherwise flat ground. They were constructed as follows (see Fig. 1). Within a square area with sides of length s, a circle with diameter $0.9s$ was centered. All depth values outside the circle (i.e. in the object base plane, which in the initial display is the same as the image plane) were set to zero. For each of three positions inside the circle (located at the vertices of an equilateral triangle), the depth was specified as either $+h$ (a distance h in front of the object base plane, closer to the observer), 0 (in the object base plane), or $-h$ (behind the object base plane). A smooth spline was constructed, using the data splining capability of the GRID3 3-dimensional plotting program (Reference Note 1), which passed through the flat surround and the vertices of the triangle. For a given set of vertices, 27 shapes were constructed in this way (see Fig. 1 for some examples).

Two different sets of vertices were used to generate shapes. These were either at the corners of a triangle pointing up (designated '*u*') or of a triangle pointing down (designated '*d*'). Thus, there were 54 possible shape designations. These are denoted by indicating the trio of positions (*u* or *d*), and then specifying for each position (in the order shown in Fig. 1), whether that position is in front of the object base plane ('+'), in the plane ('*0*'), or behind it ('-'). For example, the shape denoted by '*u+-0*' consists of a bump in the upper-central area of the display, a depression in the lower-left, and a flat area in the lower-right (see Fig. 1).[1]

Displays were generated for all combinations of the 54 shapes, three dot numerosities, and three bump heights. Thus, there were 486 possible shapes. Dot numerosities were 20, 80, and 320. Bump height, h, was $0.5s$, $0.15s$, or $0.05s$, where s is the length of a side of the square ground. The 3D perspective drawings of the shapes in Fig. 1 are for the largest bump heights.

Multi-dot displays of these shapes were generated by choosing a random sample of positions on each surface, rotating the resulting set of points about a fixed axis, and projecting them onto an image plane via parallel projection. The 3D motion was a single cycle of a sinusoidal rotation about a fixed vertical axis through the center of the object base plane, with amplitude of 25 deg and period of 30 frames. Thus, each object appeared face-forward, rotated, say, to the right until it had rotated 25 deg, reversed direction and rotated to the left until it was 25 deg to the left of its initial orientation, and then reversed direction and rotated until it was again face-forward. Two rotation directions were used, indicated as '*l*' and '*r*', corresponding to whether the left or right edge of the display comes forward initially. Equivalently, this describes the side of the observer to which the shape 'faces' in the second half of the rotation (which is usually an easier way to code the response) A full description of a display might be '*u+-0l*', for example. Given the parallel projection, simultaneous reversal of depth signs and of rotation direction yields precisely the same image sequence. Thus, '*u+-0l*' and '*u-+0r*' describe the same display.[2]

After sampling, rotation, and projection, any given frame of the display consisted of *n* points in the image plane. These points were displayed as luminance dots on a dark background. The square image extent of the displays projected to a 182 × 182 pixel area subtending 4 deg of visual angle. The displays were not windowed in any way, so the edges of the display oscillated in and out with the rotation. With the 25 deg wiggle, the display reaches a minimum of 90% of its original horizontal extent.

Displays were presented on a background that was uniformly dark (approximately .001 candelas/m^2). Dots were single pixels of approximately .65 μcandles, A trial sequence consisted of a cue spot presented for 1 sec, a 1 sec blank interval, and the stimulus sequence. The stimulus sequence was followed by a blank screen, the luminance of which was the same as the background of the stimulus. The display was run at 60 Hz noninterlaced. Each display frame was repeated four times, for an effective rate of 15 new frames per second. The duration of each 30 frame display was 2 sec.

Apparatus. Stimuli were computed in advance using a Vax 11/750 computer and stored on disk. The stimuli were displayed using an Adage RDS-3000 image display system and were displayed on a Conrac 7211C19 RGB color monitor. The stimuli appeared as white dots on a black background.

Viewing Conditions. Stimuli were viewed monocularly (with the dominant eye) through a black cloth viewing tunnel. In order to minimize absolute distance cues, there was a circular aperture slightly larger than the square display area. Stimuli were viewed from a distance of 1.6 m. After each stimulus presentation, the response was typed by the subject on a computer terminal. Room illumination was dim (illuminance was approximately 8 cd/m^2).

Procedure. Each of the 486 displays (54 shapes/rotations, three numerosities, three heights) was viewed once by each subject. The displays were presented in a mixed-list design in four sessions of 45 min. After each response, feedback was provided as to the possible correct responses. There were always two responses for each stimulus which were scored as correct (given perceptual reversals). For the flat stimuli, four possible answers were correct.

Subjects were shown perspective drawings of the shapes (as in Fig. 1), and were instructed as to how they were constructed and named. They were told that they would be shown multi-dot versions of these shapes, and would be required to name the shape displayed and its rotation direction as accurately as possible. They were told to use any method they chose to remember and apply the shape and rotation designations.

Each subject ran in several practice sessions in order to become familiar with the task and the method of response. Practice sessions consisted of half of the easiest stimuli (the 320 dot 0.5*s* height stimuli), or 27 trials. All subjects ran approximately five practice sessions, until accuracy was at least 85% correct.

RESULTS

The results of the experiment are summarized in Fig. 2. Each response was scored as correct only if both the shape and the rotation direction were correct and consistent. Thus, if ‘*u+-0l*’ was the display, responses *u+-0l* and *u-+0r* were considered correct. Any other response was incorrect. There were occasional responses with the correct shape and the incorrect rotation direction (66 such errors, 4.5% of all responses, 10% of all errors). Subjects later indicated that these were a result of difficulties with the response, rather than from a truly mis-rotating percept. Regardless, such responses were treated as incorrect.

As expected, accuracy improved both with the numerosity and with the amount of depth displayed. An ANOVA was computed treating numerosity, height, and subjects as treatments, and shapes/rotations as the experimental units. Both numerosity and degree of depth are highly significant ($p<.0001$). Subjects significantly differ from one another ($p<.0001$). The three-way interaction was significant ($p<.01$), indicating that the interaction of height and number differed among subjects (see Fig. 1). No two-way interactions were significant.

Confusion matrices were computed for each subject, pooled across the nine conditions, two rotation directions, and two possible designations of each shape (it was thus a 27 × 27 matrix). An insufficient quantity of data was collected to enable us to confidently draw specific conclusions from the error data. Table 1 is a summary of identification errors, pooled across subjects. The hypothesis that errors are distributed uniformly across the nine error classes is easily rejected ($\chi^2=1031.12$, $df=8$, $p<.001$). It appears that four types of errors were the most prevalent. Large single bumps were highly confusable, especially the distinction between ‘*d+++*’ and ‘*u+++*’, but also that between ‘*d+++*’ and ‘*d0++*’, etc. Errors were made in horizontal location of the shape within the ground (e.g. ‘*d0+0*’ was reported as being ‘*u+00*’, or ‘*d++0*’ as ‘*u+0+*’). Errors were also made in judging the width of the bumps (e.g. ‘*d+00*’ reported as ‘*u0++*’). Finally, where both a bump and a concavity were present, occasionally one of the two was not noticed. It is interesting that in every case of this type of error (the ‘Missed Smaller Feature’s and ‘Missed Equal Size Feature’s of Table 1, and the less common missed larger features), the response was of a single bump toward the observer. In other words, in the presence of a perceived convexity, a concavity is occasionally missed, but not the other way around. On the other hand, when only one nonzero depth was present (a single bump or concavity), it was very rare for subjects to give a response containing multiple depth signs.

DISCUSSION

We have introduced a new objective task for measuring the perceptual effectiveness of the kinetic depth effect: shape identification. It is a measure of perception of shape. With the current lexicon of shapes, it measures whether the subject can globally determine the areas which are in front of the ground, and which are behind. Because of the large set of shapes and the systematic way in which it was constructed, and the large set of possible responses, it is very difficult to perform this task without a global perception of shape.

For example, suppose that a subject wanted to perform this task by only measuring instantaneous velocities at a small number of spatial positions, say, halfway through the motion sequence. Clearly, measurements at six positions — the corners of both triangles used in specifying the shapes — would be sufficient, but it would be difficult for the subject to make the measurements accurately, and very difficult to determine which of 108 possible responses was consistent with them. On the other hand, less measurements do not suffice. If measurements were made at only the three corners of one of the triangles, the shape is incompletely specified. If all three measurements indicate zero depth at those positions (given the known rotation speed), there still may be a bump present created using the other triangle. Simple velocity measurement strategies do not help subjects. Too many positions must be monitored, and the cognitive load is too great.

We have previously argued (Landy, Dosher, and Sperling, 1986) that measurement of the full effect of stimulus manipulations on the KDE requires several subject responses in order fully to describe the richness of the percept. These responses included judgements of coherence (whether the multi-dot stimulus coheres as a single object), rigidity (does the object stretch?), and depth extent (what is the amount of depth perceived). These different aspects of the percept are partially coupled, but they do not all increase with the same stimulus manipulations. For example, in some subjects the addition of exaggerated polar perspective to a display increases the perceived depth extent while decreasing the sense of object rigidity.

In the current experiments, this richness of the KDE percept is not being measured. Instead, we are simply measuring to what extent the display was effective in creating a global sensation of depth, and hence of shape. Other aspects such as depth extent or rigidity are not measured. Increasing the depth extent displayed does improve performance, as we have seen, but we have not measured the depth extent perceived.

Neither have we measured the degree of rigidity perceived in the displays. In fact, nonrigid percepts were reported by subjects. One particular example was very common. Shapes with both bumps and concavities (e.g. $u{+}{+}{-}$) were occasionally seen in a nonrigid mode. Rather than seeing one area forward, another back, and the whole thing rigidly rotating, observers just as readily perceived both areas as being in front of the object ground, rotating in opposite directions (this percept looks rather like a mitten with the thumb and fingers alternately grasping and opening). This particular nonrigid percept occurred most often when the number of dots was large and the depth extent was at its largest. In this stimulus condition, with mixed-sign shapes it is clearly visible that the two bumps cross (in the rigid mode, one sees through the bump to the concavity behind it when they cross). This is an example of a failure of the 'rigidity hypothesis' (Ullman, 1979; Schwartz & Sperling, 1983; Braunstein & Andersen, 1984; Adelson, 1985), since a perfectly rigid figure is easily perceived in a non-rigid mode. These stimuli are multi-stable, with more than two possible stable percepts. In our experiments, again, we are not measuring this richness of the percept, but merely whether global shape has effectively been perceived. Subjects with the nonrigid percepts were required to compute the name of one of the possible rigid percepts that was consistent with what they perceived.

Several cues may be leading to the percept of shape in this task. One cue is dynamic change in texture density. The shapes are generated in such a manner that, face-on, the expected local dot density across the display is uniform. As the shape rotates, areas in the display becomer more dense or sparse as the areas in the shape that they portray become more or less slanted from the observer with the rotation. Theoretically, the observer could use this cue to determine the shape. In another paper (Landy, Dosher, Sperling, and Perkins, 1987) we report results of experiments in which this cue is manipulated. By varying dot lifetimes the density cue can be eliminated, keeping local average dot density constant across the display. This manipulation does not lower performance significantly, and hence the dot density cue is not necessary. By reducing dot lifetimes to a single frame, one can create a display containing only the density cue. Performance in this condition is poor, and 3D shape is not perceived. The dot density cue is thus an insufficient cue to depth in these displays.

Other possible cues relate to dot motion. Subjects could either be deriving shape from a global optic flow field (instantaneous velocity vector measurements across the field), or from measurement of relative position of various dot pairs across an extended span of time. Models of the KDE have been based on both optic flow measurement (Koenderink & van Doorn, 1986) and relative distance measurement (Ullman, 1984; Landy, 1987). By reducing dot lifetimes to two frames, one can create a display where a global optic flow field is available (although noisy), but where the span of time is minimized over which measurements can be made of the relative positions of pairs of points. It turns out that subjects are quite effective at the shape identification task with such displays (Landy et al, 1987). This may be taken as evidence against the relative position measurement models (Ullman, 1984; Hildreth & Grzywacz, 1986).

We have found that shape identification performance increases with the number of dots displayed and the extent of depth portrayed. Neither of these results is surprising. The numerosity result is consistent with previous, more subjective, measures of the depth perceived in KDE displays (Green, 1961; Braunstein, 1962). Increasing the number of dots provides the observer with more samples of the motion of the shape portrayed. Increasing depth extent increases the range of velocities used. Both manipulations increase the observer's signal-to-noise ratio in the task, where noise sources may be both external (such as position quantization in the display and poor shape sampling) and internal.

ACKNOWLEDGMENT

The work described in this paper was supported in part by a grant from the Office of Naval Research, Grant N00014-85-K-0077.

NOTES

1) Note that there are in fact only 53 distinct shapes possible, since '*u000*' and '*d000*' both designate the same shape: a flat square.

2) There are 108 possible display designations (54 shape designations and 2 rotation directions). In fact, there are only 53 possible shapes as indicated in Note 1. Given depth reversals, there are only 53 unique display types. In the experiments, we in fact use 54 different displays, including two tokens of the flat shape, which is denoted equally accurately as *u000l*, *u000r*, *d000l*, and *d000r*. Chance performance depends on subject strategy as a result, unfortunately. Repeated responses of *u000l* (and its equivalents) yields a guaranteed performance of 2 in 54 correct. Random guessing yields an expected performance of just over 1 in 54 correct.

REFERENCE NOTES

1) J. Anthony Movshon, GRID3, Version 6.3, Perspective projection of a three-dimensional surface, 1981.

REFERENCES

Adelson, A. H. (1985). Rigid objects that appear highly non-rigid. *Investigative Ophthalmology and Visual Science, Supplement*, **26**, 56.

Andersen, G. J., and Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, **34**, 356-362.

Bennett, B. M., and Hoffman, D. D. (1985). The computation of structure from fixed-axis motion: Nonrigid structures. *Biological Cybernetics*, **51**, 293-300.

Braunstein, M. L. (1962). Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology*, **64**, 415-420.

Braunstein, M. L. (1977). Perceived direction of rotation of simulated three-dimensional patterns. *Perception and Psychophysics*, **21**, 553-557.

Braunstein, M. L., and Andersen, G. J. (1981). Velocity gradients and relative depth perception. *Perception and Psychophysics*, **29**, 145-155.

Braunstein, M. L., and Andersen, G. J. (1984). A counterexample to the rigidity assumption in the visual perception of structure from motion. *Perception*, **13**, 213-217.

Clocksin, W. F. (1980). Perception of surface slant and edge labels from optical flow: a computational approach. *Perception*, **9**, 253-269.

Green, B. F., Jr. (1961). Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, **62**, 272-282.

Hildreth, E. C., and Grzywacz, N. M. (1986). The incremental recovery of structure from motion: Position vs. velocity based formulations. Proceedings of the Workshop on Motion: Representation and Analysis, IEEE Computer Society #696. Charleston, South Carolina, May 7-9, 1986.

Hoffman, D. D. (1982). Inferring local surface orientation from motion fields. *Journal of the Optical Society of America*, **72**, 888-892.

Hoffman, D. D., and Bennett, B. M. (1985). Inferring the relative three-dimensional positions of two moving points. *Journal of the Optical Society of America A*, **2**, 350-353.

Hoffman, D. D., and Flinchbaugh, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics*, **42**, 195-204.

Inada, V. K., Hildreth, E. C., Grzywacz, N. M., and Adelson, E. H. (1986). The perceptual buildup of three-dimensional structure from motion. *Investigative Ophthalmology and Visual Science, Supplement*, **27**, 142.

Koenderink, J. J., and van Doorn, A. J. (1986). Depth and shape from differential perspective in the presence of bending deformations. *Journal of the Optical Society of America, A*, **3**, 242-249.

Landy, M. S. (1987). A parallel model of the kinetic depth effect using local computations. *Journal of the Optical Society of America A*, in press.

Landy, M. S., Dosher, B. A., and Sperling, G. (1986). Assessing kinetic depth in multi-dot displays. *Bulletin of the Psychonomic Society*, **19**, 23. or Kinetic depth judgements of multi-dot displays. (1987). etc. etc.

Landy, M. S., Dosher, B. A., Sperling, G., and Perkins, M. E. (1987). The kinetic depth effect and models for visual motion detection???, etc.

Lappin, J. S., Doner, J. F., and Kottas, B. L. (1980) Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, **209**, 717-719.

Longuet-Higgins, H. C., and Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London, Series B*, **208**, 385-397.

Petersik, J. T. (1979). Three-dimensional object constancy: Coherence of a simulated rotating sphere in noise. *Perception & Psychophysics*, **25**, 328-335.

Petersik, J. T. (1980). The effects of spatial and temporal factors on the perception of stroboscopic rotation simulations. *Perception*, **9**, 271-283.

Proffitt, D. R., Bertenthal, B. I., and Roberts, R. J. (1984). The role of occlusion in reducting multistability in moving point-light displays. *Perception & Psychophysics*, **36**, 315-323.

Schwartz, B. J., and Sperling, G. (1983). Luminance controls the perceived 3-D structure of dynamic 2-D displays. *Bulletin of the Psychonomic Society*, **21**, 456-458.

Todd, J. T. (1982). Visual information about rigid and nonrigid motion: A geometric analysis. *Journal of Experimental Psychology: Human Perception and Performance*, **8**, 238-252.

Todd, J. T. (1984a). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception and Psychophysics*, **36**, 97-103.

Todd, J. T. (1984b). The analysis of three-dimensional structure from moving images. In D. Ingle, M. Jeannerod, and D. Lee (Eds.), *Brain Mechanisms and Spatial Vision*. The Hague: Martinus Nijhoff, 1984.

Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.

Ullman, S. (1984). Maximizing rigidity: the incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception*, **13**, 255-274.

Wallach, H., and O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, **45**, 205-217.

Webb, J. A., and Aggarwal, J. K. (1981). Visually interpreting the motion of objects in space. *Computer*, **14**, 40-49.

Figure and Table Legends

Figure 1) Shapes, rotations, and their designations. In the experiment, subjects were required to name the shape and rotation direction perceived. Shapes were smooth splines of a flat ground and three points which were either toward the observer ('+'), neutral or in the flat ground ('*0*'), or away from the observer ('-'). These three points were at the corners of one of two equilateral triangles: (a) with the odd point up ('*u*'), or (b) with the odd point down ('*d*'). The numbers specify the order in which the three point's depth sign are to be reported in designating the shape. (c) The various combinations result in a lexicon of 53 shapes. Typical examples are illustrated here as perspective plots. The orientation of these plots relative to the viewing direction is indicated on the first example. (d) Two motions were simulated. Both were sinusoidal rotations about a vertical axis through the center of the object ground. The object either first rotated to face the subject's right, then to the subject's left, then returned face-forward ('*l*'), or in the opposite direction ('*r*').

Figure 2) Performance on the task as number of points in the simulated shape was varied. The parameter is the height of the bumps. Performance increased with both numerosity and bump height.

Table 1) Summary of identification errors, pooled across subjects, bump heights, numerosities, rotation directions, and depth reversals. The first column gives a description of eight common error types, along with a miscellaneous category. If a bump and a depression were present in the display, and only one of the two was indicated by the subject, this was called a 'Missed Feature Error'. If the bump and depression are of equal extent on the base plane (e.g. '*u+-0*'), then this is called a 'Missed Equal Size Feature'. If they are of unequal extent, and the smaller of the two is not reported, this is categorized as a 'Missed Smaller Feature'. Any display containing only one depth sign (such as '*u+00*') reported as containing both depth signs (e.g. '*u0+-*') is categorized as an 'Add a Depth Sign' error. For a given row in the table, the second column presents examples of errors of that type. The third column lists the number of cells in the confusion matrix which correspond to an error of a given type, while the fourth column provides the total number of errors in all cells of that type. The last column is the average number of errors in cells of that type. For comparison, the bottom row of the table provides summary information. In particular, there were 0.8 errors per cell overall.

Description	**Examples**	**Number of Cells**	**Number of Errors**	**Ratio**
Large bumps	$u+++$ vs $d+++$	2	29	14.5
Horizontal Extent	$u0++$ vs $d+00$	4	34	8.5
Missed Smaller Features	$u++-$ reported as $u++0$	6	30	5.0
Diagonal to Large Bump	$u++0$ reported as $u+++$ or $d+++$	8	23	2.9
Missed Equal Size Feature	$u+0-$ reported as $u+00$	12	29	2.4
Diagonal Extent	$u++-$ reported as $u+0-$	8	16	2.0
Small Horizontal Location Error	$u+00$ vs $d0+0$	16	26	1.6
Add a Depth Sign	$u+00$ reported as $u+-0$	168	39	0.2
Other Errors		478	360	0.8
All Errors		702	586	0.8

Table 1

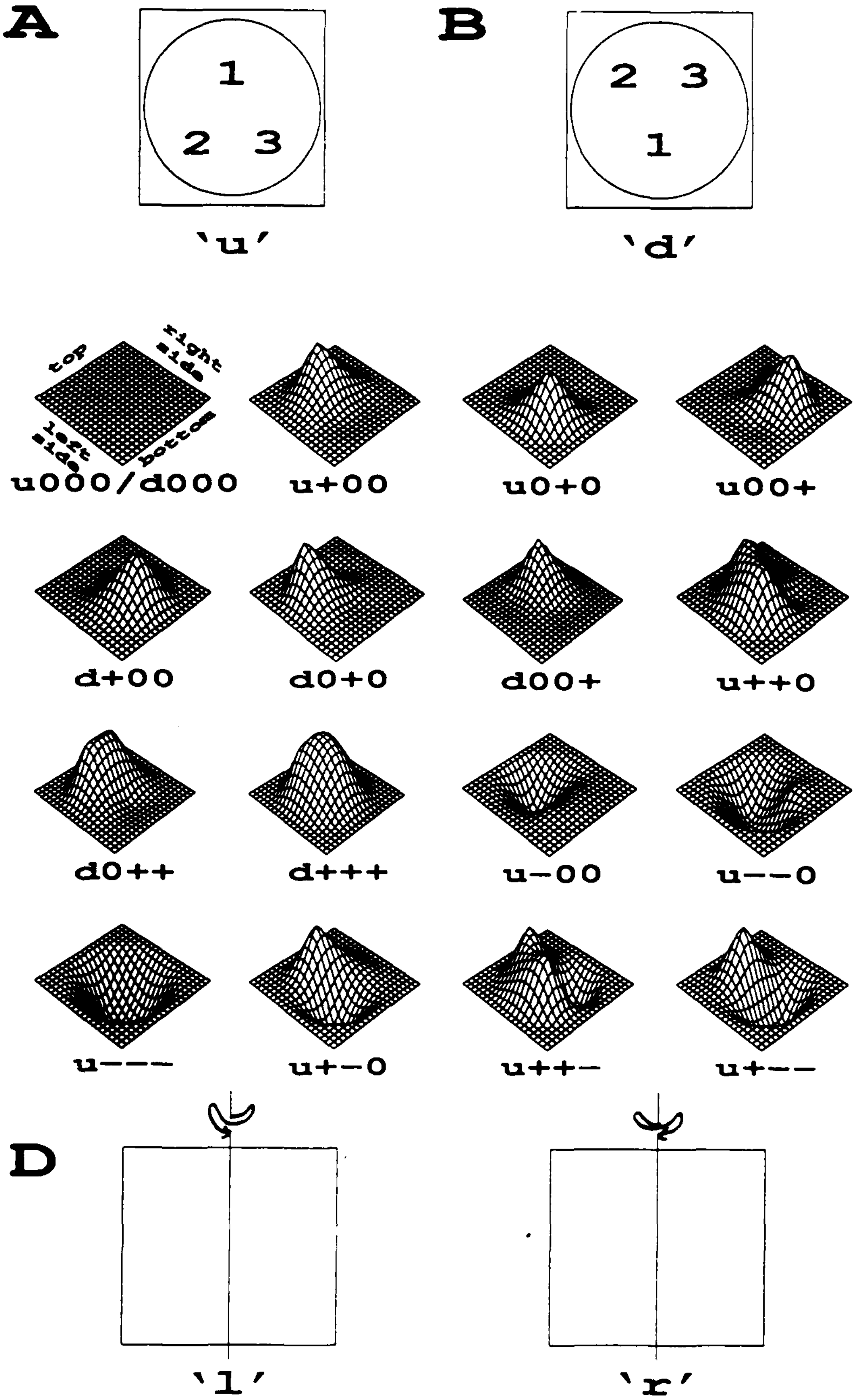
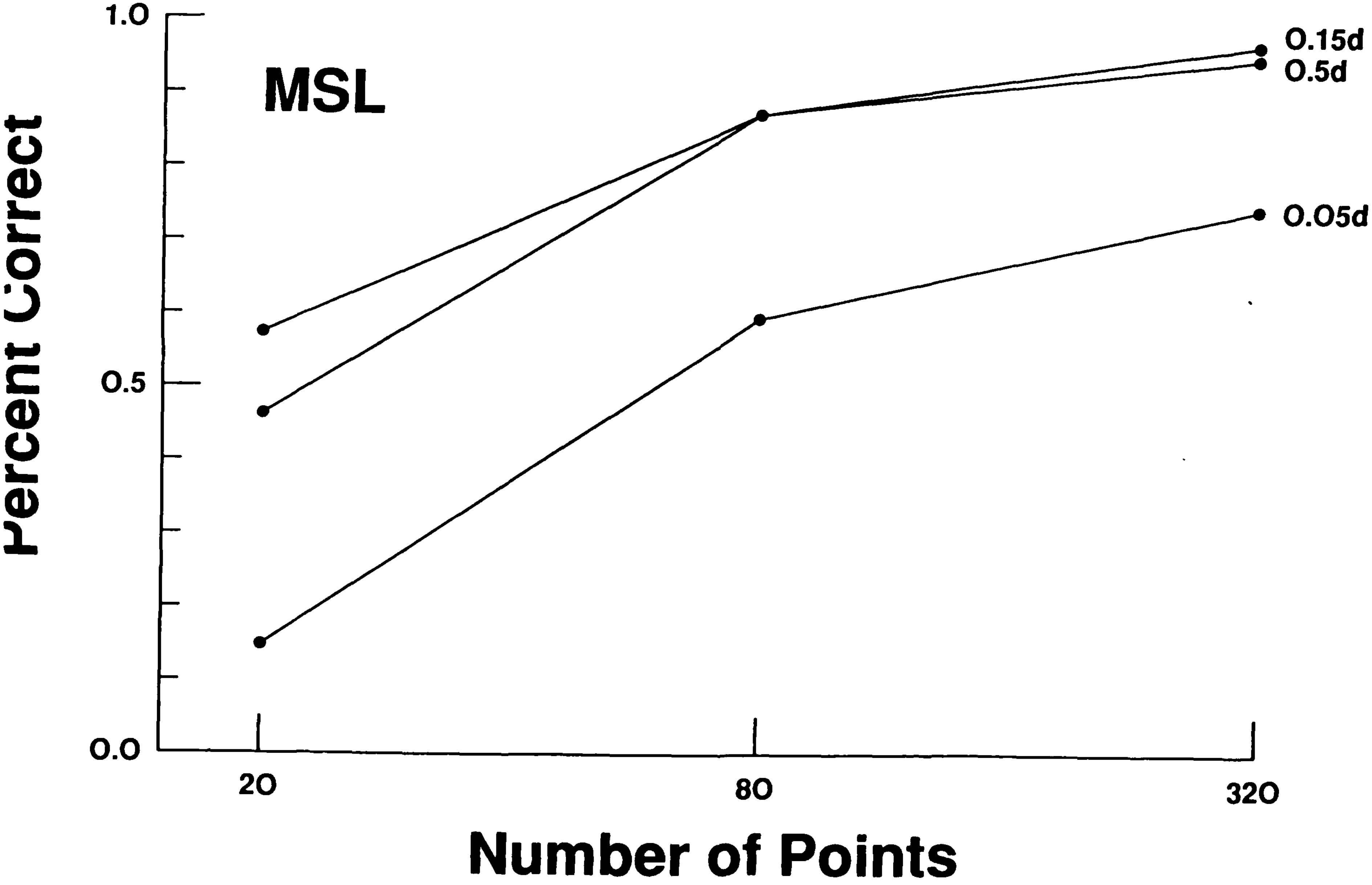
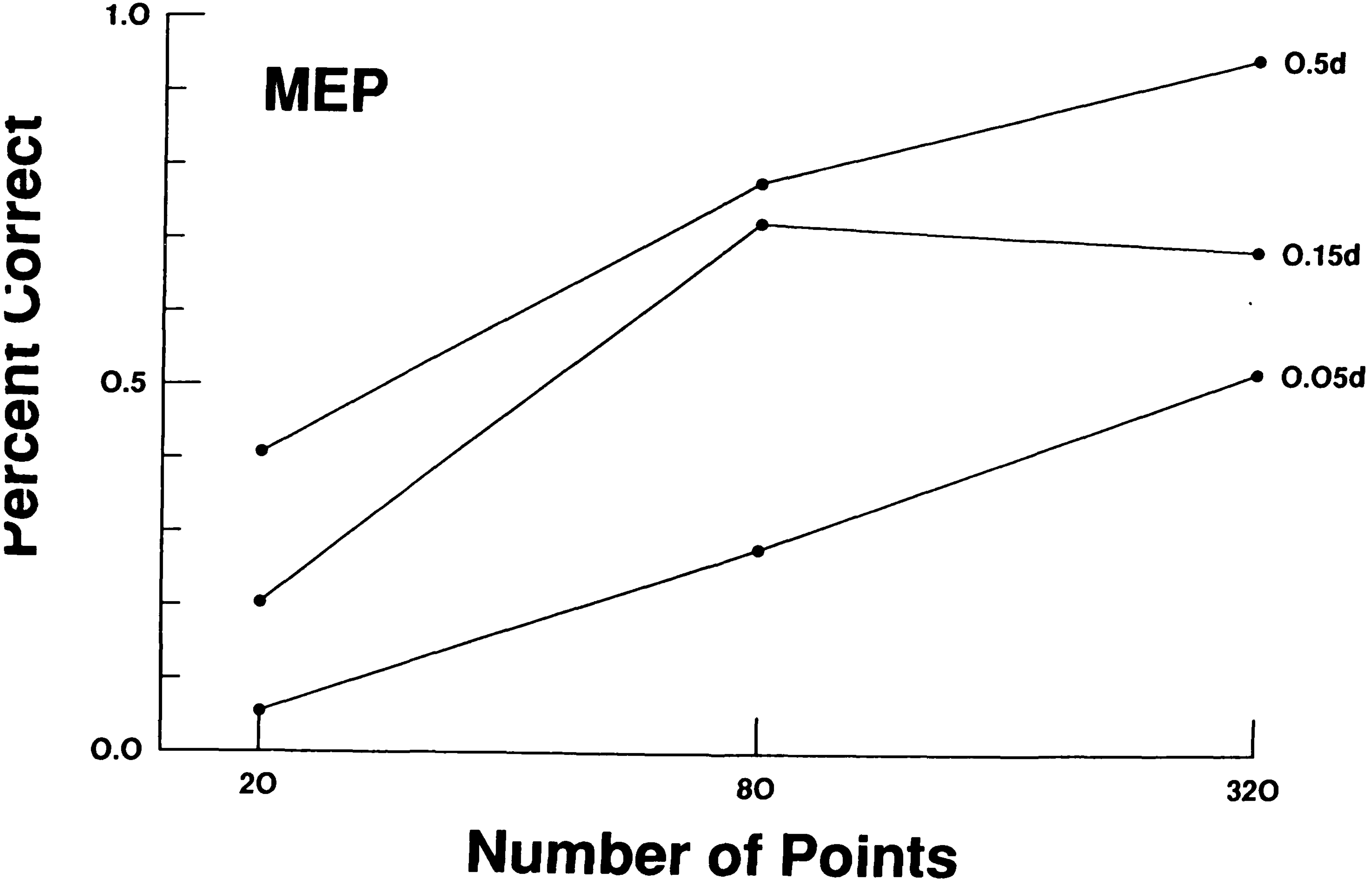
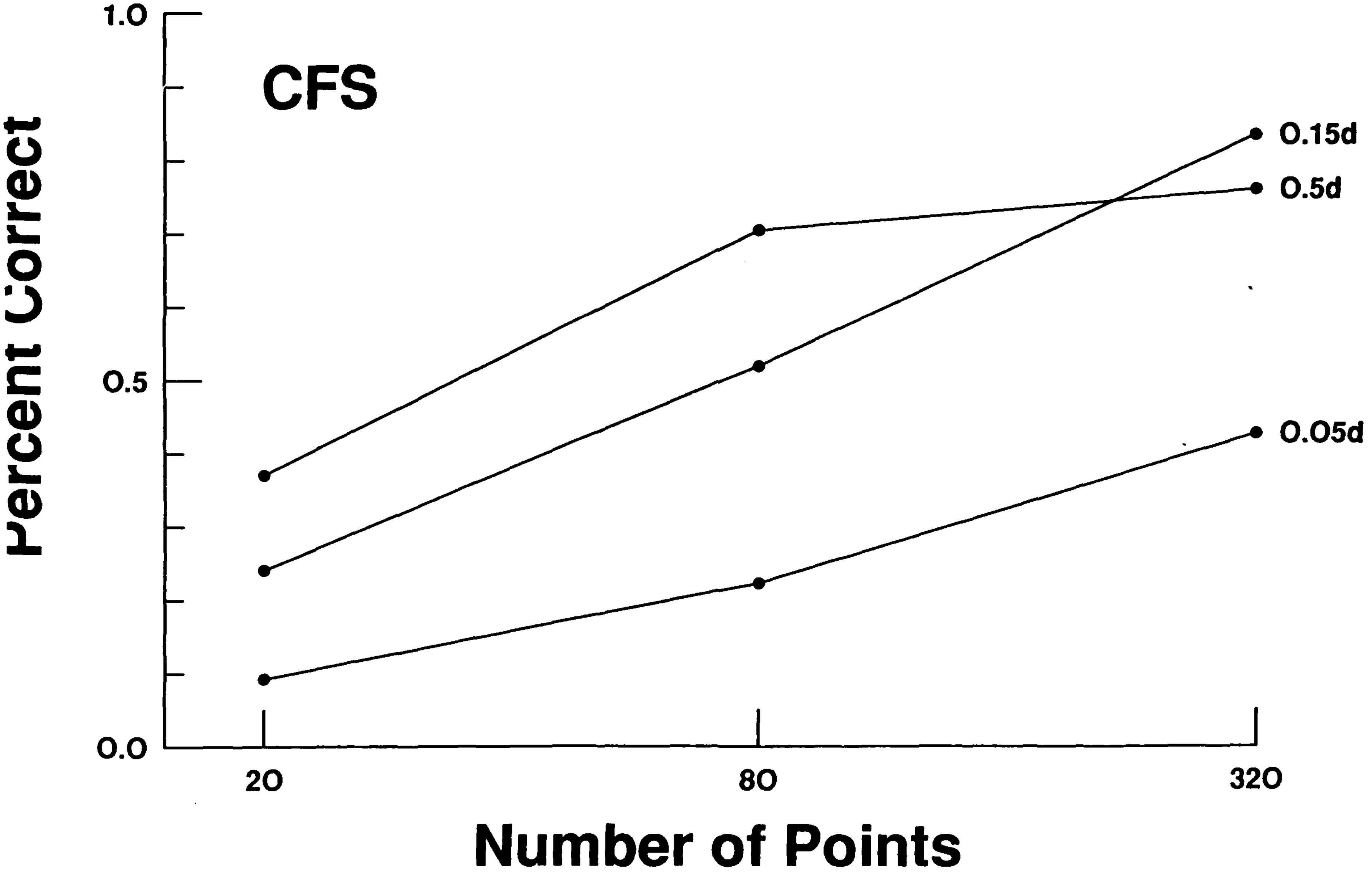

A
1
2 3
'u'
B
2 3
1
'd'
C
top
right side
left side
bottom
u000/d000
u+00
u0+0
u00+
d+00
d0+0
d00+
u++0
d0++
d+++
u-00
u--0
u---
u+-0
u++-
u+--
D
'l'
'r'

figure 2


MSL
Percent Correct
1.0
0.5
0.0
20
80
320
Number of Points
0.15d
0.5d
0.05d


MEP
Percent Correct
1.0
0.5
0.0
20
80
320
Number of Points
0.5d
0.15d
0.05d


CFS
Percent Correct
1.0
0.5
0.0
Number of Points
20
80
320
0.15d
0.5d
0.05d

END

7-87

DTIC