Perception, 1975, volume 4, pages 173-178

Learning to see complex random-dot stereograms

John P Frisby, Jeremy L Clatworthy
Department of Psychology, University of Sheffield, Sheffield S10 2TN, England
Received 9 August 1974

Abstract. Many observers of complex random-dot stereograms find that the depth effect takes several seconds, or even minutes, to develop. Julesz (1971) has noted that giving a priori information to such observers about the nature of the 'hidden' cyclopean object appears to facilitate their stereopsis. An experiment is reported which investigated this possible facilitation. Naive subjects were shown a complex stereogram following various kinds of preliminary assistance, ranging from simply telling them about the amount of depth they could expect to see to showing them a full-scale model of the cyclopean object. Surprisingly, no benefit from such assistance could be demonstrated. All observers improved their stereopsis perception times with repeated presentations of the stereogram, showing that they could, in principle, benefit from assistance. A follow-up study three weeks later revealed that a substantial part of this improvement was maintained, indicating that the perceptual learning involved can last for a considerable period of time.

1 Introduction

Julesz (1964, 1965) has demonstrated that 'stereopsis perception time' for a random-dot stereogram can be as short as 50 ms if the stereogram is a simple one containing just two planar surfaces with small disparity (e.g. a square standing out in depth from a surround by virtue of a 7′ visual angle disparity). He contrasts this brief perception time with the much longer ones required for complex random-dot stereograms containing many depth planes with large disparities, e.g. those portraying such intricate cyclopean objects as spiral staircases, domes, intersecting ellipsoids, etc (Julesz 1971, p 201). His explanation for the latter lengthy times is that complex stereograms require the observer to make a series of vergence movements in order to bring corresponding dots in the two fields into Panum's fusional area, a necessary precursor to binocular fusion. This explanation can also account for the fact that repeated viewings of a given complex stereogram result in a reduced perception time: it could be that with practice the observer learns to make an efficient series of vergence movements—he learns to avoid divergent movements when convergent ones are appropriate and vice versa. Thus Julesz (1971, p 216-217) has likened the process of fusing a complex stereogram to exploring a maze, with wrong 'turnings' (i.e. wrong vergence movements) prolonging perception time.

This eye movement explanation can also make sense of an intriguing informal observation made by Julesz (1971, p 201). He claims that perception times can be speeded up considerably if the observer is given an a priori knowledge of the 'hidden' cyclopean object. Even simply a verbal description of what he should be seeing can apparently help. He suggests that such advance knowledge helps shift attention (and thus presumably vergence eye movements) to appropriate depth planes and by so doing it greatly diminishes search time for fusing all the dots. The experiment reported here sought to confirm Julesz's informal observation in controlled laboratory conditions and to discover what sort of a priori knowledge is most helpful. In particular, this experiment investigated whether or not a priori knowledge simply about the amount of depth to expect is as helpful as knowledge about the actual shape of the cyclopean object. Accordingly, the perception times of naive observers were measured for seeing a complex random-dot stereogram following various kinds of preliminary assistance. Knowledge about the amount of depth to expect was conveyed to some subjects just by telling them verbally and to others by providing a fixation point which was suitably positioned in depth to reflect the greatest depth to be seen in the stereogram. Knowledge about the shape of the cyclopean object was given to other subjects either verbally or by providing monocularly visible contours in each half stereogram. Knowledge about both depth and shape was given to yet other subjects either verbally or by showing them a full-scale model of the cyclopean object. Finally, a control group of subjects was given no special help whatsoever.

2 Method

2.1 *Subjects*

103 first-year university students of psychology, 68 women and 35 men, took part. No subject had any previous experience of random-dot stereograms.

2.2 *Apparatus and stimulus*

A complex random-dot stereogram depicting a spiral staircase (Julesz 1971, p 122) was used as the stimulus throughout. It was back-projected on to a ground glass screen which was viewed by the subject through suitably crossed polarising filters mounted in a headrest set 114 cm from the surface of the screen. The images of the two halves of the stereo-pair were so aligned that the base of the spiral lay in the depth plane of the screen. The tip of the spiral had a convergent disparity of $1^\circ\ 7'$ with respect to its base so that when it was binocularly fused it seemed to protrude about 30 cm from the screen. The overall size of the stereogram was $15^\circ \times 15^\circ$. The white parts of the spiral surface had a brightness of $34 \cdot 3$ cd m^{-2} and the dots had a brightness of $<3 \cdot 4$ cd m^{-2} (as measured with an SEI spot photometer).

2.3 *Experimental conditions and procedure*

The experimental session, which lasted about 10 min, began with the subject reading the following formal instructions:
"This is an experiment on depth perception. You will be shown a number of stimuli and each time your task will be to press a button as soon as you see an object-in-depth. I will be able to tell from your button press how long it has taken for the depth effect to develop. You will first of all be shown a picture of a fly in 'depth' and in 'non-depth' to familiarise you with the nature of the task. Please wear spectacles if you would normally do so for the distances involved".

The subject was then shown the fly vectograph which forms part of the Wirt (or Titmus) Stereotest [1]. This vectograph contains a stereo photograph of a magnified housefly (approximately 12 cm x 6 cm) which stands out in depth from the page of the test booklet by about 5 cm when a subject with normal stereovision views the vectograph wearing polaroid spectacles. The subject was asked to inspect the vectograph with and without the spectacles so that he became reasonably familiar with the meaning of the terms 'depth' and 'non-depth' as used in the instructions. A few minutes were spent informally discussing this distinction. It was hoped that this preliminary training would enable the subject to know the kind of experience for which he should press the button, yet at the same time keep him naive with respect to random-dot stereograms.

The subject was then given a further set of instructions to read. These instructions differed for the various experimental conditions but all included the following remarks: "The stimuli proper will appear on the screen in front of you. I will warn you when the stimuli are about to appear by saying 'ready'. Please look at the screen carefully

[1] Available from Clement Clarke International Ltd, Instrument Division, 16 Wigmore Street, London W1H 0DH.

following this warning. Note that sometimes the depth effect might take some time to develop. After you have pressed the button, describe the object which you see".

There were seven experimental conditions and each subject was randomly allocated to one of them given the constraint that approximately equal numbers of subjects performed in each condition. The various conditions and the distinctive sentences which characterised their instructions were as follows:

(a) *No help.* "Be sure to press the button as soon as you see the object-in-depth".

(b) *Full verbal prompt (spiral and depth).* "The object-in-depth which you are to look for is a *spiral* coming out from the screen towards you. The spiral will protrude about *one foot* from the screen surface. Be sure to press the button as soon as you see the spiral in depth".

(c) *Partial verbal prompt (spiral).* "The object-in-depth which your are to look for is a *spiral.* Be sure to press the button as soon as you see the spiral in depth".

(d) *Partial verbal prompt (depth).* "The object-in depth which you are to look for will seem to protrude about *one foot* from the screen surface. Be sure to press the button as soon as you see the object-in-depth".

(e) *Model.* "The object-in-depth which you are to look for is depicted in a model which I will show you. Note that it is a *spiral* which protrudes about *one foot* from the screen surface. Be sure to press the button as soon as you see the spiral in depth". Subjects in this condition were then shown a full-scale replica of the spiral held up against the screen.

(f) *Fixation point.* "Stare at the black dot throughout. Be sure to press the button as soon as you see the object-in-depth". In this condition a small black dot was arranged (via optical superimposition) to appear just where the tip of the spiral lay once binocular fusion had taken place.

(g) *Monocular contours.* The instructions used in this condition were the same as those used for conditions (b) *full verbal prompt.* However, a version of the basic spiral stereogram was used which had monocularly discriminable contours traced on to it which followed the outline of the spiral (see figure 1).

The first stimulus presentation was given as soon as the instructions had been read, any questions arising from them answered and the 'ready' warning announced. Following his button press, the subject was asked to describe the object to which he had responded, and if he failed to give an adequate verbal description of the shape of the object (descriptions such as 'mound' or 'rings' were unacceptable), the direction

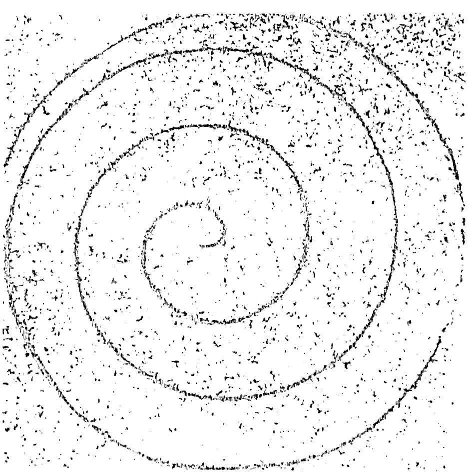

Figure 1. Spiral stereogram used in condition (g) *monocular contours.*

of the depth (towards or away from him) or the amount of the depth (around 1 foot) was noted down. The stimulus was then removed from view and a further five stimulus presentations given, each preceded by a brief rest of about 20 s and a warning call of 'ready'. The purpose of these trials was to observe the nature and course of the perceptual learning in each condition. The spiral replica was shown only on the first trial of condition (e) *model,* as it became redundant once the subject had seen the spiral itself. The further five trials in condition (g) *monocular contours* employed the uncontoured spiral stereogram used in all the other conditions, so that the course of learning in each condition could be observed in comparable circumstances as far as the stimulus was concerned.

Following the six presentations of the spiral, the subject's stereoacuity was assessed by means of the circles subtest of the Wirt Stereotest. This subtest presents the subject with a series of nine sets of four circles in vectograph form. One circle in each set of four possesses a certain amount of disparity which causes it to 'float' above the page of the test booklet for an observer with normal stereovision. The subject has to state in a forced choice manner which circle in each set of four is the one floating out in depth. The nine sets of circles present increasingly fine disparities.

The subject's final task was to complete a brief questionnaire which asked him whether he had seen a stimulus similar to the spiral before, whether he had ever suffered from an eye problem such as squint and whether he had discussed the experiment with previous subjects. The purpose of including the test of stereoacuity and the questionnaire was to eliminate from the analysis any subject who had defective stereovision or previous knowledge of the experiment and/or random-dot stereograms.

3 Results

Six subjects were eliminated from the data analysis because their descriptions of the object-in-depth following the first stimulus presentation clearly revealed that they had responded on the wrong basis. Three of these had delayed responding until they could think of a suitable verbal label to describe what they were seeing; the other

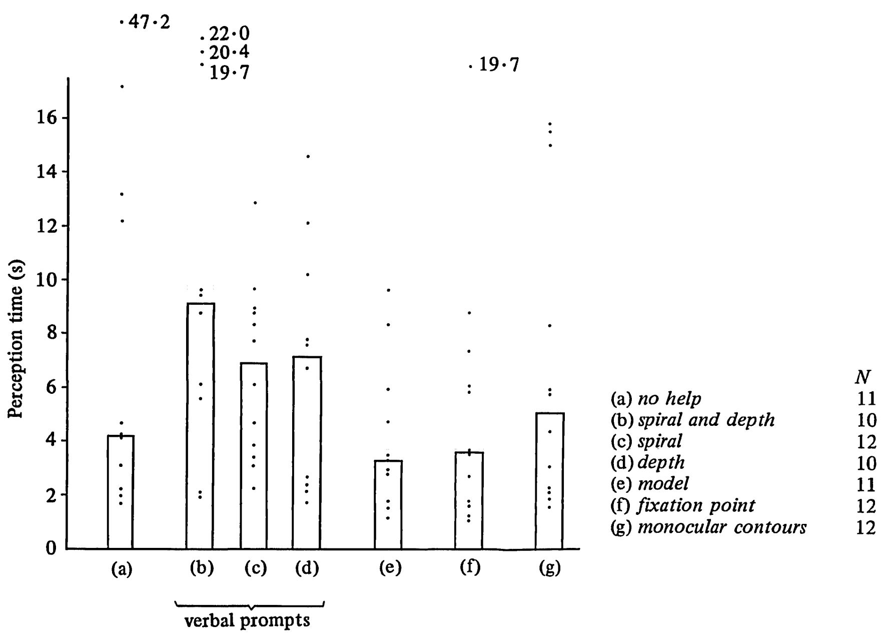

Figure 2. Stereopsis perception times for the first presentation of the spiral stereogram. The dots show individual scores; the columns show group medians.

three responded too soon, i.e. before the spiral had emerged. A further 19 subjects were eliminated because they failed to give 100% performance on the stereoacuity test. All subjects who reported a history of squint except one fell into this latter group. The exception not only had good stereoacuity but he also produced one of the fastest perception times for the spiral so there seemed little reason to exclude his data from the analysis. These various eliminations left a final total of 78 subjects, 49 women and 29 men.

Figure 2 presents both the group median and individual subject perception times for seeing the spiral on the first presentation. The differences between the group medians are nonsignificant (Kruskal–Wallis H statistic = 7·86, d.f. = 6). Figure 3 provides the group median times for all six presentations and it clearly reveals that all groups showed a learning effect (similar to the one reported by Ramachandran and Braddick 1973) such that all groups rapidly achieved perception times of around 1 s or better.

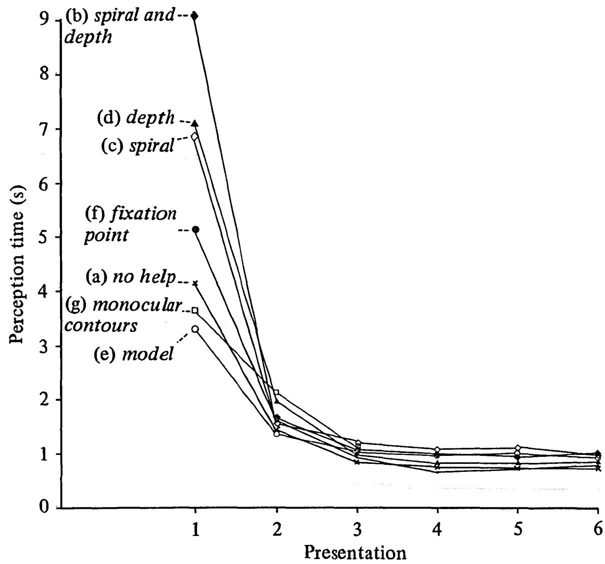

Figure 3. Group median perception times for each of the six presentations of the stereogram in each condition.

4 Discussion

The absence of any marked differences between the groups is very surprising. The only sign that the *no help* subjects [condition (a)] were any worse off than the others is that they produced the longest individual perception time (47·2 s). Equally, the only sign that the condition of greatest assistance, the (e) *model,* was any easier is the fact that it produced no particularly long times (see figure 2). This absence of prominent differences between conditions is even more curious considering that all groups, and indeed all subjects, showed a clear learning effect over presentations (see figure 3). Thus all groups were, in principle, capable of benefitting from some kind of assistance. And yet, as figure 1 clearly shows, the various kinds of assistance had little, if any, influence on the perception times recorded.

This result is so counter-intuitive that it forces a re-examination of the design of the experiment and its various controls. One possible criticism is that the fly vectograph may not have given subjects an adequate basis for making a response. Were there, for instance, many subjects who delayed making a response until they were quite sure that the 'thing' they were seeing in fact constituted a suitable 'object-in-depth'? It will be remembered that several subjects were rejected for response difficulties of this kind. Could there have been many more who escaped elimination despite the careful questioning to screen them out? We find this unlikely, but, even so, a factor of this kind can hardly explain the similar median times observed in conditions (a) *no help* and (e) *model.* The spiral replica must surely have given a very clear idea of what was required for a satisfactory response.

It would be wrong, of course, to conclude from this experiment that the perception of complex random-dot stereograms cannot be facilitated by a priori knowledge after all. What can be concluded is that it is surprisingly difficult to demonstrate such facilitation in naive observers should it in fact exist[2]. Thus it could be that the experiment included the wrong kind of assistance—although it is difficult to imagine a more helpful condition than (e), the *model,* whose spiral replica even had its surface dotted to mimic the texture of the random dot spiral itself. Perhaps it could be that a priori knowledge can be of benefit but only to relatively experienced observers who know how to take advantage of it.

Table 1. Mean, median, and range stereopsis times for the first and last presentations of the spiral in the main experiment, and the single presentation of the follow-up study. The 41 subjects who provided the data for this analysis included no members of group (g) *monocular contours*, thus ensuring that all subjects received the same spiral stereogram as stimulus throughout.

	Mean	Median	Range
First presentation	8·9	8·3	1·5–47·2
Last presentation	0·9	0·8	0·4–2·0
Follow-up presentation	5·9	3·4	1·3–22·5

Finally, it is worth reporting the results of a small follow-up investigation. It proved possible to run 41 of the 78 subjects on a single presentation of the same spiral stereogram about three weeks after the conclusion of the main experiment. The same general procedure was used, with each subject reading a set of formal instructions which told them to expect the same spiral stimulus as before and reminding them to respond by pressing the button as soon as the spiral emerged in depth. Table 1 presents the mean, median, and range times for these subjects on the first and last presentations of the spiral in the main experiment and on the single presentation in the follow-up. No subject achieved in the follow-up his previous asymptotic level of performance. On the other hand, it was evident that some perceptual learning had been maintained in that follow-up times were significantly faster than times for the first presentation of the main study. (Wilcoxen's T statistic $= 193$, $N = 41$, $p < 0{\cdot}005$, two-tailed). This supports Julesz's claim (1971, p 217) that the perceptual learning involved in seeing a complex random-dot stereogram can be retained for a considerable time.

References

Julesz B, 1964 "Binocular depth perception without familiarity cues" *Science* **145** 356–362

Julesz B, 1965 "Texture and visual perception" *Scientific American* 212 38–48

Julesz B, 1971 *Foundations of Cyclopean Perception* (Chicago: University of Chicago Press)

Ramachandran V S, Braddick O, 1973 "Orientation-specific learning in stereopsis" *Perception* 2 371–376

Saye A, Frisby J P, 1975 "The role of monocularly conspicuous features in facilitating stereopsis from random-dot stereograms" *Perception* 4 159–171

(2) Saye and Frisby (1975) have, however, demonstrated that monocular contours can facilitate stereopsis for naive observers of a simple two-planar random-dot stereogram with very large disparity.