Proc. Int. Symposium Teleoperation and Control, 249-260
July 1988, © IFS Ltd and authors, ISBN 1-85423-009-3

The effects of display flicker rates on task performance

P M Schweiwiller, V M Reading
Institute of Ophthalmology, UK
A A Dumbreck and E Abel
Harwell Laboratory, UK

ABSTRACT

Flicker in a CRT display can upset the control of saccadic eye movement, even at frequencies above flicker fusion. The effect decreases with increasing field rate, therefore operator performance of a demanding visual task may be facilitated by using a display with a high scan rate. The experiment described here compared the speed and accuracy with which a visual search task was performed when the display field rate was 50, 80 and 100Hz.

INTRODUCTION

CRT based displays are widespread in remote viewing applications. Most use field rates in the region of 50 to 75 Hz and will often appear to flicker appreciably.

Flicker can be minimised by careful choice of viewing conditions and display design [1], however there is evidence that the human visual system responds to the intermittency of illumination even when no flicker is visible. Tyler recorded electrical activity of the brain synchronous with a display field scan at frequencies above flicker fusion [2]. Brindley demonstrated the perception of beats between electrical and visual stimulation of the eye at frequencies where either stimulus alone would be invisible [3]. It is possible that these responses have undesirable effects and may be responsible for the commonly reported finding of a slower reading speed from a CRT than from typescript [4].

Conventional video signals can be converted by digital signal processing to a format with a higher field rate immediately prior to display [5]. Thus completely flicker-free viewing can be obtained, circumventing the problems of equipment compatibility. With the falling cost of digital field storage this will become an economic proposition in the near future. If there is also an effect on operator performance, such equipment will be particularly advantageous for surveillance and remote viewing by television.

Wilkins showed that flicker can cause the enlargement of saccadic eye movements [6]. The extent of the enlargement depends on the display field rate and is evident when flicker is not noticeable. At 50Hz saccades can be 11% larger than at 100Hz. This may necessitate corrective saccades, which themselves might be enlarged. A practical implication may be that accurate fixation of the eyes will be facilitated at a high field rate. A demanding visual search or inspection task, involving many changes of fixation, will be performed more quickly and possibly more accurately than at a lower field rate. If the minimum time between saccades is 100msec [7] then every ten saccades requiring correction will add 1 second to the search time and produce a measurable difference in performance.

To test this hypothesis, we compared observer performance of a visual search task using display field rates of 50,80 and 100Hz.

VISUAL SEARCH TASK

Subjects were required to find and count the number of circular targets hidden in a background pattern of random dots. There were 100 tasks and each contained between 1 and 20 targets. (See figure 1a for an example.) For each number of targets (1 to 20) there were five patterns, generated by a microcomputer and stored on disc prior to the experiment. The disposition of the targets was randomised over the screen, but evenly distributed between the four quadrants of the screen and none were less than two targets-widths distant from any other. The resolution of the pattern was 256 by 192 pixels and the targets were drawn in cells 8 by 8 pixels. The background pattern of random dots had approximately equal numbers of black and white dots and was the same for all 100 tasks.

APPARATUS

The experiment was controlled and timed by a BBC Master Series microcomputer, which retrieved the appropriate display pattern from the disc and transferred it as a bit-image, to a display controller, based on an Acorn ATOM. The ATOM had been modified in several ways and was able to generate video signals at field rates nominally of 50,80 and 100Hz.

The picture was 262 lines, non-interlaced of which 242 lines were active. The display had a bright border 25 lines wide at the top and bottom, and the equivalent of 58 pixels at either side. This was unavoidable in the graphics mode, but under the viewing conditions used the border did not flicker noticeably, even at 50 Hz.

Two display monitors were used; a BARCO CD351 to display 50 and 80Hz and a MANITRON VLR20101 for 50 and 100Hz. The BARCO was a colour monitor with a dot-pitch of 0.32mm and P22 phosphor; the MANITRON a high quality black and white with P4 phosphor. Both of these are medium-short persistence phosphors [8] and the overall decay waveforms, when the colour components have been weighted according to the CIE photopic luminous efficiency curve [9], are very similar for each monitor.

Room Lighting, Viewing Distance and Picture Size

The viewing conditions used were chosen to minimise the visibility of the picture line structure which was rather coarse, to make the task suitably difficult and the picture flicker at 50Hz imperceptible.

Room lighting was subdued throughout and entirely from overhead incandescent lamps. The illuminance falling on a horizontal surface was 170 Lux where the subject was seated and there were no obvious reflections from the monitor screen; glare was minimal.

The BARCO was viewed from 160cm and the MANITRON from 150cm. The active picture area (including the border) subtended 13 by 10 degrees of visual angle, and a target 0.3 by 0.3 degrees. The viewing distances were slightly different so that each monitor subtended the same visual angle. (See Figure 1b.)

Monitor Luminance

Monitor brightness and contrast were matched as accurately as possible, using a spot photometer. The black squares of a chequered test pattern were set (nominally) to 3 Cd/m sq and the white ones to 55 Cd/m sq, giving a contrast ratio of 18.3:1.

Display Resolution

Since the sharpness of the display could affect search times, the resolution was measured.

The microcomputer was used to generate linear gratings of different spatial frequencies up to the maximum where alternate pixels along a line were on. As a measure of system response, the luminance modulation of the display was measured at each frequency. With a close-up lens attached to the spot photometer it was possible to measure the luminance of a spot 0.4mm in diameter, so the highest frequency grating could be encompassed. Modulation is defined as:-

$$M = (L_w - L_b) / (L_w + L_b) \tag{1}$$

Where L_w and L_b are the luminance of the white bars and black bars respectively.

Although modulation was found to fall off slightly at the highest spatial frequency for field rates of 80 and 100Hz, the difference was not noticeable. Comparing the measured modulations with modulation sensitivity functions of the human visual system obtained by Van Meeteren and Vos [8] for a field, 17 by 11 degrees suggests that the picture degradation would have to be about 200 times greater before the pattern became invisible.

SUBJECTS

Twelve subjects, aged between 22 and 36 years and naive as to the purpose of the experiment, participated; there were 10 per sub-experiment and eight were common to both parts. None were frequent VDU users. Within the last three years each had been ophthalmically screened and had 6/9 or better visual acuity.

EXPERIMENTAL DESIGN

The experiment was, of necessity, in two parts; the first compared subject performance of the search task at field rates of 50 and 80Hz, using the BARCO monitor, whilst the second, compared 50 and 100Hz using the MANITRON. It was expected that the two 50Hz conditions would give a sufficiently similar baseline for the comparison of 80 and 100Hz to be legitimately drawn.

In each sub-experiment the subjects were divided at random into two groups (5 per group). A subject attended two sessions and completed all 100 search tasks in each. The first group saw the higher field rate in their first session and the lower rate in their second, whilst the second group saw the lower field rate first, and the higher rate second.

The two subjects new to the second part of the experiment were placed in different groups.

EXPERIMENTAL PROCEDURE

A session was started by depressing a key on the microcomputer keyboard. The instructions were displayed, page by page, on the monitor as black characters on a bright background. The subjects were instructed to count the targets as quickly and as accurately as possible. Next, five practise trials were presented for familiarisation, but from which no data were recorded.

Presentation of the 100 search tasks followed. The order of tasks was always the same - a pseudo-random sequence such that no consecutive tasks had the same number of targets. A trial commenced with the screen blanked. After about 3 seconds the display was revealed and the subject began to count the number of targets. When he had an answer he depressed the keyboard space bar and the screen was again blanked. The subject entered his answer.

For a correct answer no feedback was given; the screen was cleared and the next task presented. However, if the answer was incorrect the subject was informed and the task repeated. A maximum of three attempts was allowed for each task (although the subjects did not know this) and if the third attempt was still incorrect the experiment proceeded as if the answer had been correct. The experimenter recorded any erroneous entries which were nevertheless valid answers (eg. 9 instead of 19).

The computer recorded for each task, the number of attempts taken, the time for which the task had been displayed and, right or wrong, the subject's answer. At the end of a session the subject was asked whether he had noticed flicker in the display and for any comments. After the second session he was also asked which session had been the easier and whether he thought that the picture had been any different than before.

RESULTS, ANALYSIS AND DISCUSSION

Most of the subjects took about 50 minutes for a session, but the best performers took 30 mins, and the slowest around 75 minutes.

In drawing any conclusions from the comparison of response times it is important to take into account the error rate (or indeed the response times when considering error rates). It is possible that subjects trade-off accuracy against a fast performance or vice-versa and therefore a quantitative comparison of response times where the error rate does not remain constant can be misleading since two qualitatively different performances are being compared (slow and accurate versus fast but inaccurate).

Error Rate

The error data were analysed with a two-way analysis of variance of the number of errors made, with field rate and the number of targets. There was no statistically significant difference in the total number of errors between any of the field rates but as expected, the number of targets was significant ($F(19,3920)=12.31$; $p<0.0001$). There was no significant interaction between these two factors. It is therefore legitimate to compare subject performance by comparing response times. Figure 2 shows the number of errors made for each number of target in part one (50 vs 80Hz with the BARCO monitor). Similarly figure 3 show the number of errors in the second part.

Response Times

Only the response time for the first attempt at a task has been used. It is possible that spurious results might have been introduced using the second and third attempts, where subjects may have reasoned their answer could not have been wrong by more than one or two and simply adjusted their previous answer.

The response times were examined by regressing mean response time on the number of targets presented. Linear regression was used for both ad hoc and a priori reasons. First, because it is a simple model which fitted the data well (typically explaining about 85% of the variance), and second because response time might reasonably be expected to vary in an approximately linear way with number of targets presented. For each trial, there will be a component of response time due to searching the display, a second due to "registering" each target (to fixate and count a target) and a third due to pressing the space bar. Whilst it is difficult to know exactly what cognitive processes are involved with each component, it is reasonable to suppose that the first and third will be largely independent of the number of targets and that the second will be related proportionally. The hypothesised effect of picture flicker could affect either or both of the first two components and hence be manifest as a change of gradient or both intercept and gradient.

If the residual variance, ie. the variance not explained by the regression, was significantly reduced by partitioning the data in some way (e.g. by field rate) and fitting a linear regression, in turn, to the data on each side of the partition, then these linear regressions would be significantly different.

Figures 4 and 5 show respectively the results for the first and second parts of the experiment. Each point is the mean of 5 trial by 10 subjects and the error bars are plus or minus one standard error of the mean.

The following differences were found:-

Monitor			Difference	p value
BARCO 50Hz	vs.	BARCO 80Hz	-	n.s.
MANITRON 50Hz	vs.	MANITRON 100Hz	50Hz faster	0.0065
BARCO 50Hz	vs.	MANITRON 50Hz	MANITRON faster	<0.0001
BARCO 80Hz	vs.	MANITRON 100Hz	MANITRON faster	<0.0001

Table 1.
Regressions of Mean Response Time.

The gradient of the regression for the MANITRON at 50Hz did not differ significantly from that at 100Hz , whilst the other two significant comparisons differed in slope. ($F(1,1996)=16.11$, $p=0.0001$, and $F(1,1996)=8.66$, $p=0.0033$, respectively).

The differences between the regressions are reflected in the overall mean response times (10 subjects; 100 tasks):-

Monitor	Mean Response Time (Sec) ± Standard Deviation
BARCO 50HZ	10.67 ± 6.78
BARCO 80Hz	11.09 ± 6.45
MANITRON 50Hz	8.89 ± 4.58
MANITRON 100Hz	9.45 ± 5.07

Table 2
Mean Response Times.

It is evident that the difference between the two field rates of the MANITRON monitor was quite small and, by comparison with the difference between the two 50Hz conditions, is probably of no practical consequence.

Possible Effect of Practice

The difference between the two parts of the experiment was unexpected and because of the experimental design, it is not possible to say whether this is due to physical differences between the monitors, or can be entirely explained by a practice effect. Eight of the subjects were common to both parts, and had already completed two sessions before starting the second part.

(a) Error Rate. For these eight subjects the error rate was not constant across the sessions ($F(3,3120)=5.13$, $p=0.001$.) A Newman-Keuls analysis showed that for session 4 the rate was no different from session 1, sessions 2 and 3 were not significantly different from each other but differed from sessions 1 and 4. Session 3 and 4 are respectively, the first and second sessions of part 2.

(b) Response Times. The regressions of mean response time for these eight subjects are plotted session by session in figure 6. The ordering of the lines is consistent with a practice effect; the line for each session lies below that of the previous session and their disposition may well be enough to explain the difference between parts 1 and 2.

Since the error rate changes from session to session a quantitative comparison of the regressions is dubious, but with practice it might be expected that subjects will perform more quickly, more accurately, or both at each successive session. For the first three sessions this was true:- session 2 was faster and more accurate than session 1; session 3 was faster than session 2 for the same accuracy, but for session 4 there was a different speed-accuracy trade-off; session 4 was the fastest of the 4 sessions, but only as accurate as session 1.

Session	Mean Response Time (sec) ± Standard Deviation	No. of Errors
1	11.93 ± 6.79	111
2	10.33 ± 5.88	72
3	9.51 ± 4.89	75
4	8.65 ± 4.87	107

Table 3
Performance of Subjects Common to both Parts.

This suggests that practice explains the difference between the first three sessions and probably the difference between the two parts of the experiment, but after three sessions has little further effect.

(c) New Subjects. The two new subjects might be expected to have performed in a similar way to the ten subjects in part 1.

The higher field rate argues that these two subjects performed in a similar manner to those in part 1, but the lower field rate that they perform like those in part 2, so no conclusion can be drawn from this. Session by session, they started off slower and less accurately than the rest in part 2, but by session 2 appear to have performed at the same speed but more accurately than the other 8. This supports a practice effect, but these two subjects learnt quickly and appear to have been particularly good at the task.

Subject Comments

The majority of subjects judged the sessions to be of equal difficulty or that the second session was easier, regardless of field rate. None spontaneously reported seeing flicker, and when prompted, only one reported flicker at 50Hz, but still judged the sessions to be of equal difficulty. There was no evidence that the higher field rates were preferred.

CONCLUSIONS

The hypothesised effect of display flicker is not confirmed; there is no evidence under these experimental conditions that subjects performed a demanding visual search task with increased speed or accuracy at display field rates higher than 50Hz.

It is possible that the effects of flicker reported by Wilkins [6] are only pronounced for the highly ordered, predictable eye movements used in reading and not the more erratic ones likely to be found in the search tasks used here.

The difference between the two 50Hz displays is most plausibly explained by a practice effect rather than any physical characteristics of the monitors. By comparison, the difference between 50 and 100Hz is small and of little practical consequence.

Although the comparisons were not rigorous, the subjects did not consider any field rate easier than another and on the whole noticed little difference between any of the display conditions.

If visible flicker is not a problem, there is probably no advantage to be gained from using displays with an increased field rate for remote viewing applications.

REFERENCES

[1] Scheiwiller PM, Reading VM, Dumbreck AA, Abel E. "Flicker in Television Displays: A Review." Harwell Report AERE G 4604, Harwell Laboratory, UK, 1988.

[2] Tyler CW. "Human Brain Responses to Flicker in Visual Displays." In: Reading VM. ed. Vision and Visual Displays., Institute of Ophthalmology Course Documentation, London, 1981.

[3] Brindley GS. "Beats Produced by Simultaneous Stimulation of the Human Eye with Intermittent Light and Intermittent or Alternating Electric Current." J.Physiology (London), Vol.164, pp.157-167.(1962)

[4] Gould JD, Grischkowsky N. "Doing the Same Work with Hard Copy and with Computer Terminals." Human Factors, Vol.26, No.3, pp.323-337 (June 1984).

[5] Mehrgardt S. "Flicker-Free TV Set Design." New Electronics, 17 March 1987.

[6] Wilkins AJ. "Intermittent Illumination from Visual Display Units and Fluorescent Lighting Affect Movements of the Eyes Across Text." Human Factors,

Vol.28, No.1, pp.75-81 (February 1986).

[7] Stark L, Vossius G, Young LR. MIT Research Laboratory of Electronics, Cambridge, Mass. Quarterly Progress Report No.62, July 1962.

[8] JEDEC. "Characteristics of JEDEC registered Phosphors." JEDEC Publication No.16C., Electronic Industries Association, Washington.

[9] Hunt RW, Darby PJ. "The Measurement of Light." In: IBA Technical Review 22, Light and Colour Principles. IBA, London, 1984.

[10] van Meeteren A, Vos JJ. "Resolution and Contrast Sensitivity at Low Luminances." Vision Research, Vol.12, pp.825-833, (1972).

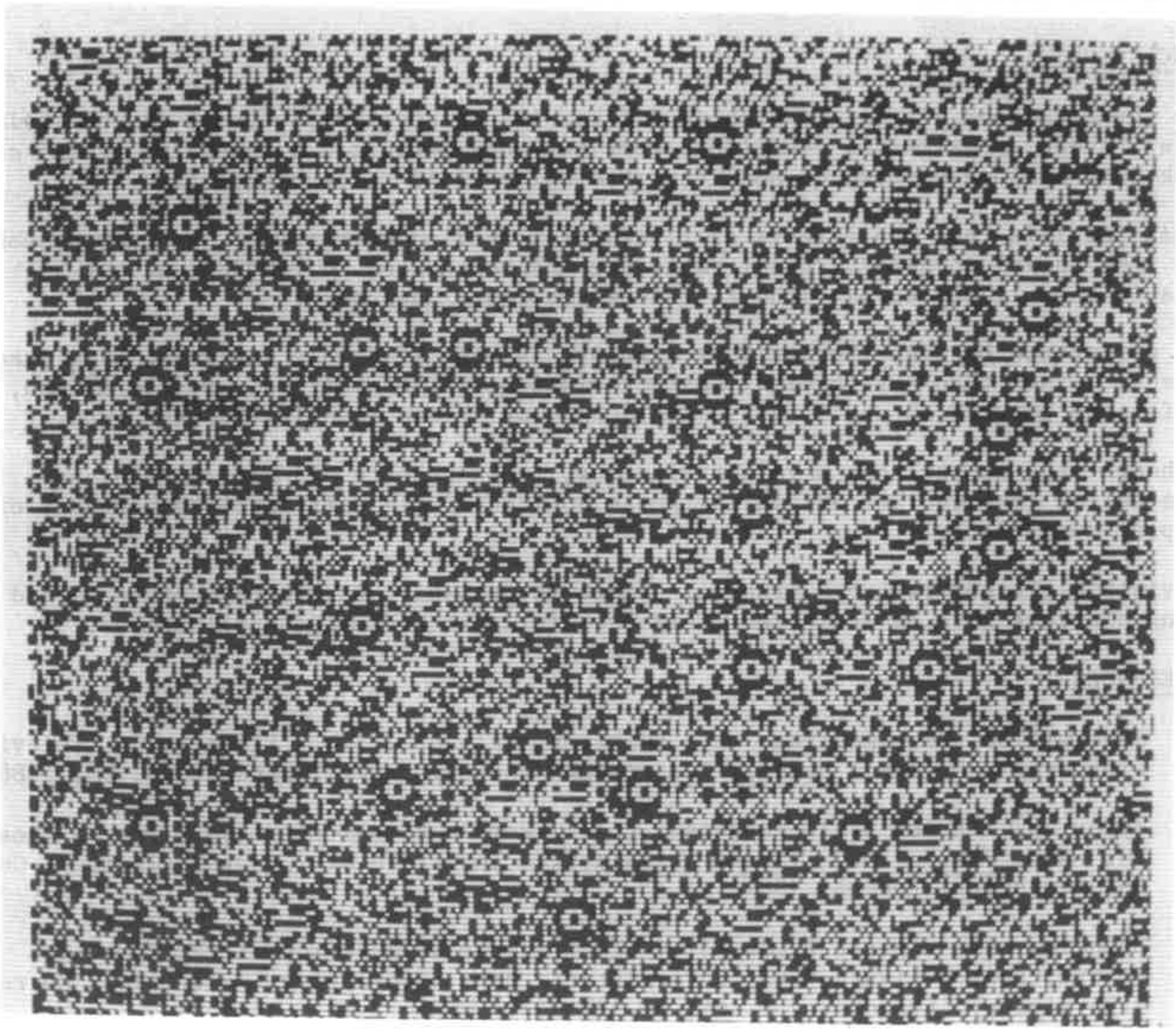

Figure 1a.
A Typical Search Task.

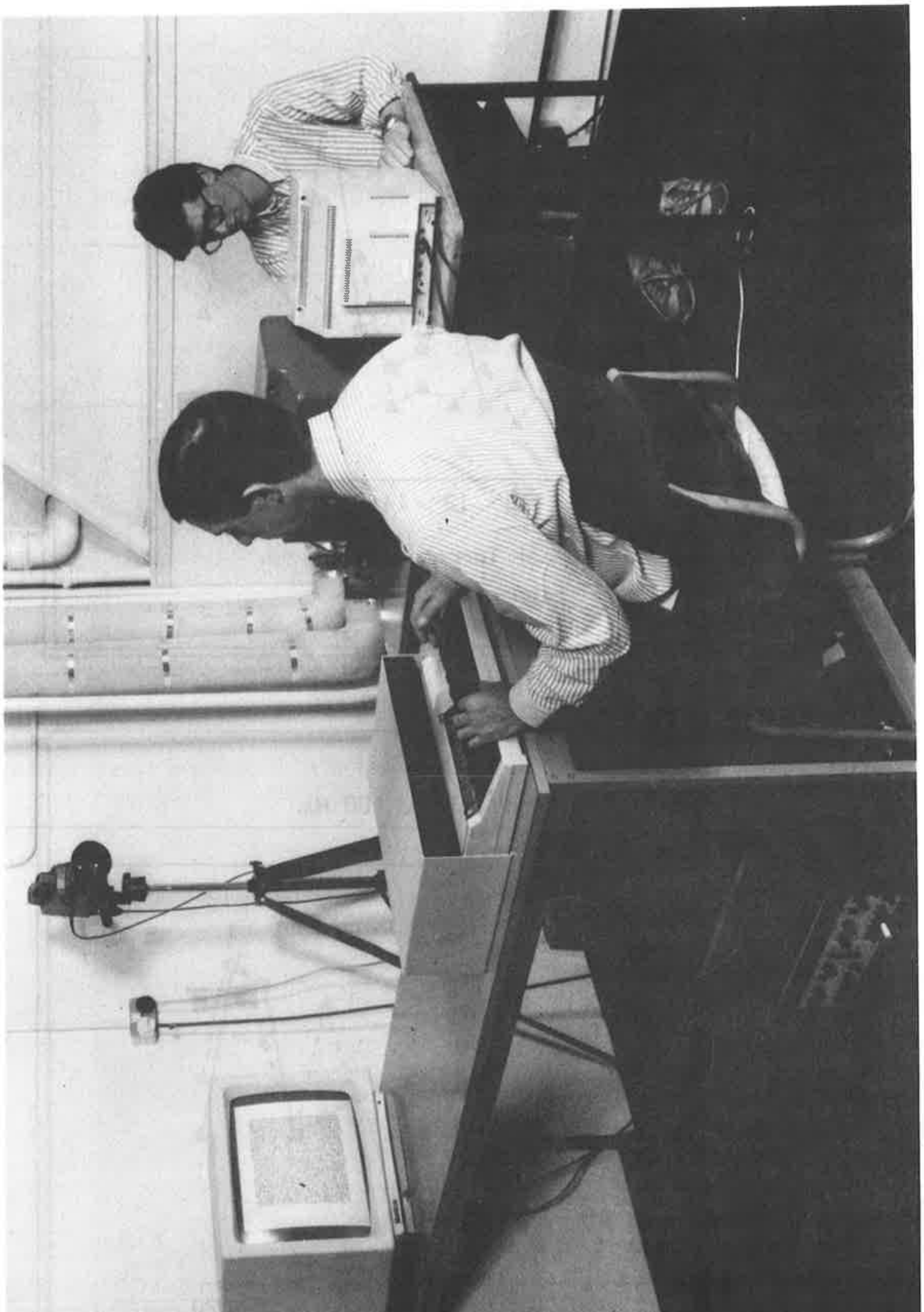

Figure 1b.
Layout of the Experiment.

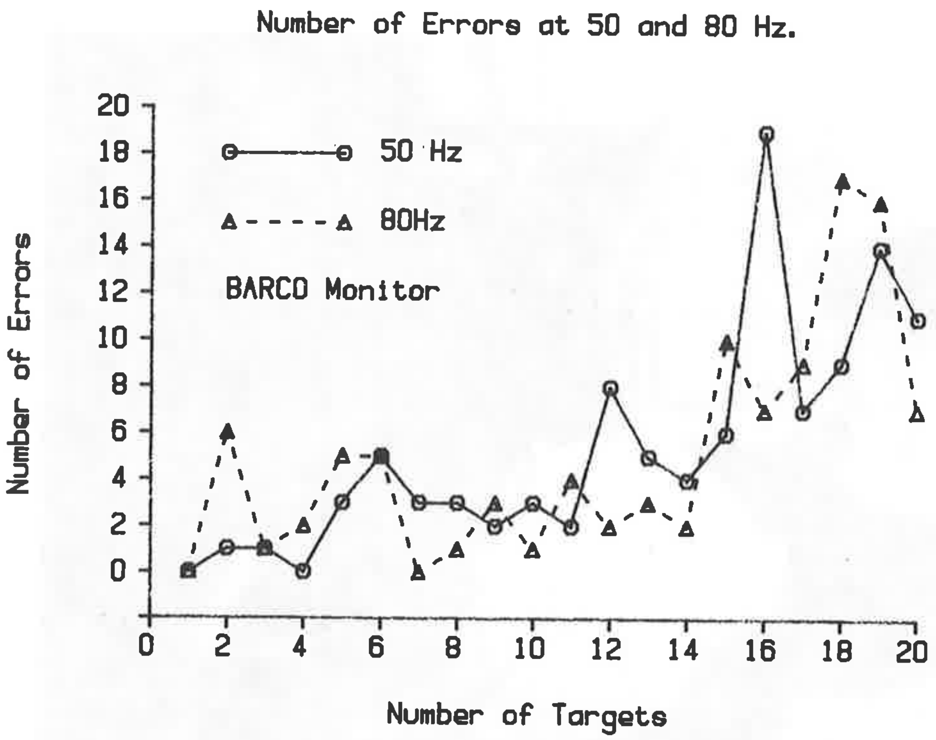

Figure 2.

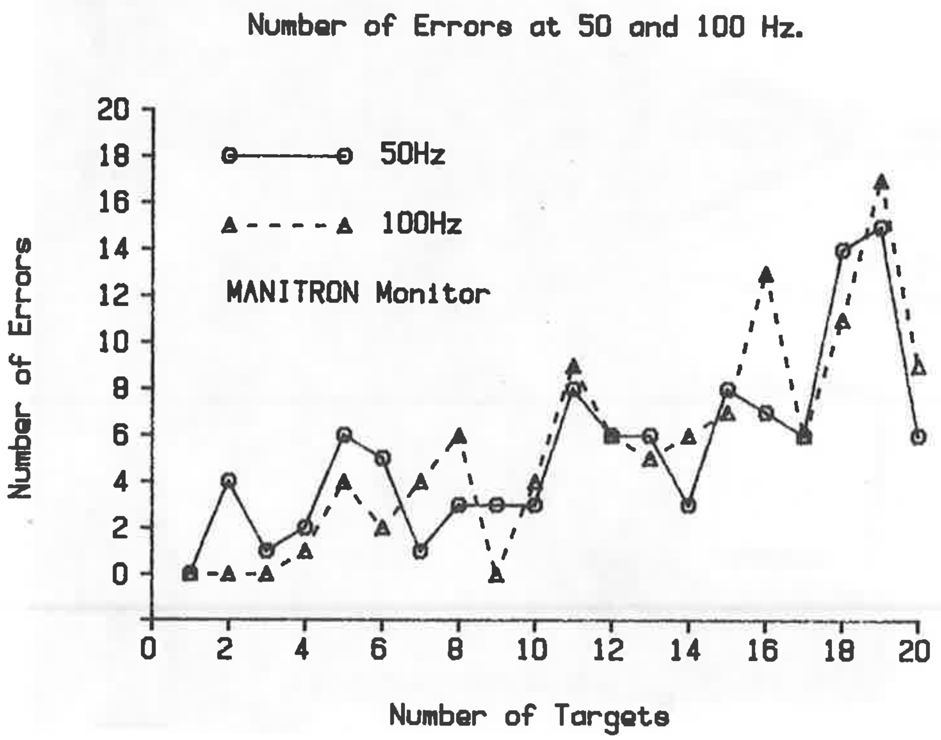

Figure 3.

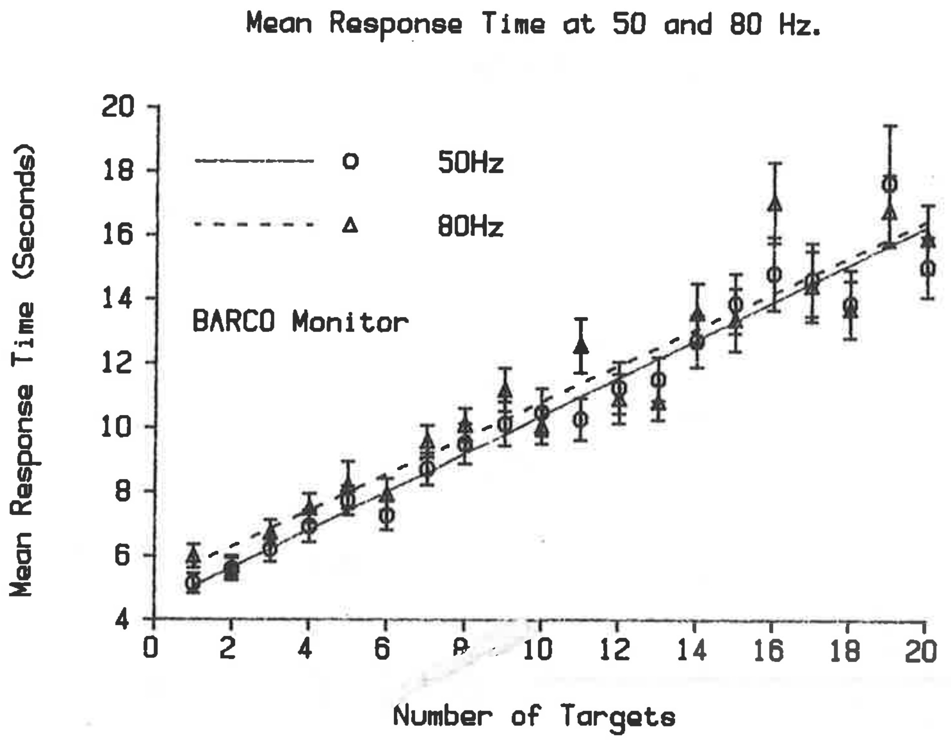

Figure 4.

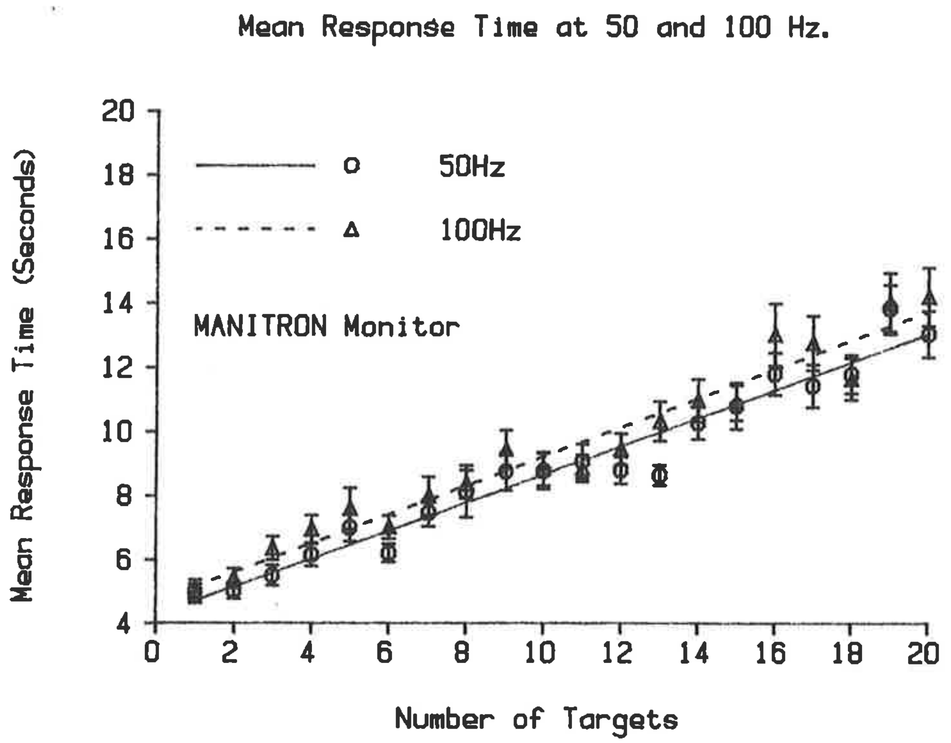

Figure 5.

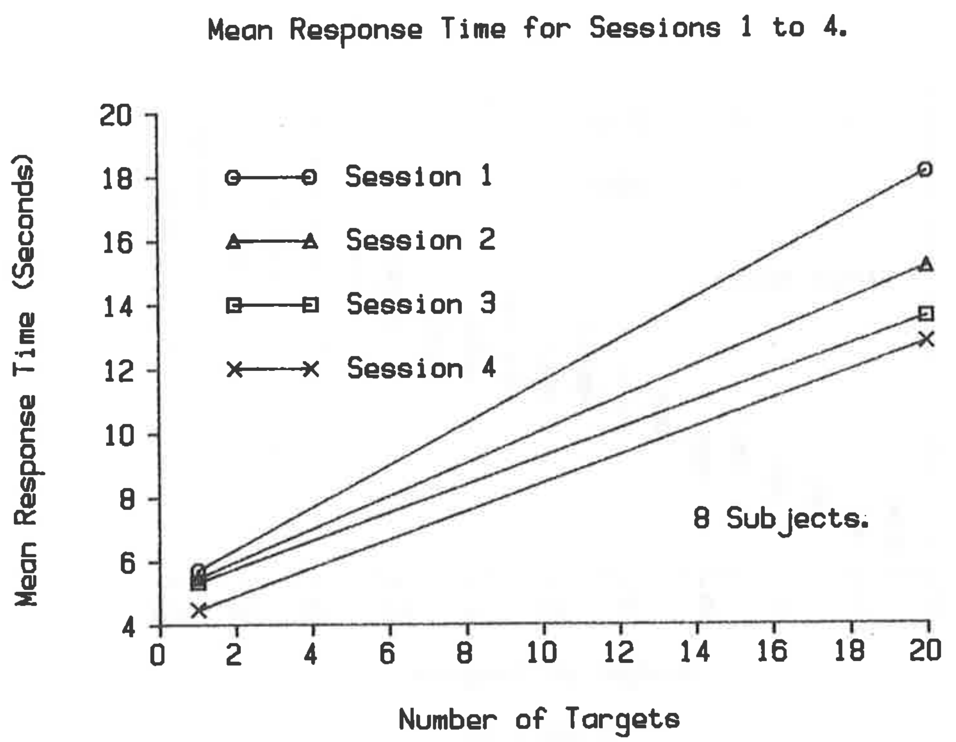
Mean Response Time for Sessions 1 to 4.
Mean Response Time (Seconds)
20
18
16
14
12
10
8
6
4
Session 1
Session 2
Session 3
Session 4
8 Subjects.
0
2
4
6
8
10
12
14
16
18
20
Number of Targets

Figure 6.