Br. J. Psychol. (1973), **64**, 3, *pp.* 383–389
Printed in Great Britain

THE EFFECTS OF MUSIC AND TASK DIFFICULTY ON PERFORMANCE AT A VISUAL VIGILANCE TASK

By D. R. DAVIES,* LESLEY LANG and V. J. SHACKLETON
Department of Applied Psychology,
University of Wales Institute of Science and Technology, Cardiff

Forty subjects performed a visual vigilance task in music or noise and when the task was difficult or easy. Response measures taken were: correct detections, commission errors, detection latencies and d' values. For the difficult version of the task a significant increase in detection latencies was found which music prevented. Broadly similar findings were obtained for correct detections. The results are compared with other studies of detection latency and task difficulty and are discussed in terms of arousal.

Much research has been concerned with factors affecting efficiency in vigilance situations and several different response measures have been used as indices of vigilance performance (Davies & Tune, 1970). These include correct detections, false positives, detection latencies and measures derived from signal detection theory such as d' and β (Green & Swets, 1966). Research has also been directed towards specifying the effects of environmental variables, such as heat, noise and varied auditory stimulation, upon monitoring performance. However, the effects of environmental variables have usually been assessed in terms of one response measure, generally correct detections, and it has not therefore been possible to compare the effects of a particular environmental variable upon different response measures. This has been an unfortunate tendency, since it is quite probable that different response measures reflect different processes, not all of which are influenced by the environmental variable in question. It may be noted, however, that some investigators (e.g. Buck, 1966) have argued that two commonly used indicators of vigilance performance, correct detections and detection latencies, both reflect changes in a single process of 'perceptual vigilance'. One aim of the present study, therefore, is to examine the way in which the introduction of an environmental variable, in this case music, affects different indicators of vigilance performance.

Comparatively few studies have examined the effects of varied auditory stimulation (VAS) or music upon vigilance performance. The available evidence suggests, however, that performance is generally improved by the introduction of music or VAS, and this can take the form of increasing the detection rate (McGrath, 1963; Tarrière & Wisner, 1962; Ware *et al.*, 1964), preventing an increase in detection latencies (Wokoun, 1963, 1969) or reducing the commission error rate (Davies *et al.*, 1969). However, McGrath (1963) also found that if his task situation was altered to allow possible distraction effects to appear, through reductions both in stimulus duration and in inter-stimulus interval, then music did not improve performance; instead it tended to make it worse. Presumably, changing the task in this way made it more

* Present address: Department of Applied Psychology, University of Aston in Birmingham, Gosta Green, Birmingham.

difficult and the present study also seeks to determine the effects of music and task difficulty (defined here in terms of the brightness difference between signals and non-signals) upon vigilance performance.

METHOD

Subjects

The subjects were 40 undergraduates (20 men, 20 women) between the ages of 18 and 28 years, who were paid for their services. Ten subjects (five men, five women) were assigned to each of the four conditions of the experiment: Noise Easy (NE), Noise Difficult (ND), Music Easy (ME) and Music Difficult (MD).

Apparatus and procedure

Subjects were tested individually and in isolation, the experimenter remaining in a separate room. Their task was to detect and respond to changes in the brightness of a light situated at eye height at a viewing distance of 1·38 m. The light, which was placed behind a filter, came on at 5 sec. intervals, its onset and offset being controlled by a Crouzet punched-card programmer, Type 227 C. The duration of the light was 1 sec. The standard brightness of the light in both the 'easy' and the 'difficult' conditions was 0·5 ft.-c. In the 'easy' conditions the signal to be detected was a brightness increment of 0·9 ft.-c., while in the 'difficult' conditions the brightness increment constituting the signal was 0·5 ft.-c. The task lasted for 40 min. and during this time 48 signals were presented, 12 in each 10 min. period. Signals were presented at apparently random times in each 10 min. period and the same cycle of signal presentation was preserved in each period of the task.

Subjects were seated behind a small table on which were placed two response keys, one labelled 'yes' and the other 'no', since the apparatus was also being used for other research concerned with detection latencies associated with different types of response in a vigilance situation. The task was thus to some extent atypical, since it is usual in vigilance situations for subjects to have only one response key. However, Whittenburg *et al.* (1956) found that using two response keys, one for responding to signals and the other for non-signal stimuli, did not produce results that were markedly different from those obtained with one response key. Subjects were instructed to press a key each time that a light appeared, whether standard or signal, but to respond with the 'yes' key whenever they thought a signal had occurred and with the 'no' key whenever they thought a standard light had been presented. Both the accuracy of each response and its latency were recorded on the output from a Kienzle printer adapted to a Venner timer, Type TSA 6614.

In the 'music' conditions a tape-recording of music consisting of 20 different musical selections was relayed to the subject (at an average intensity of 75 dbA as measured by a Dawe sound-level meter) via a Ferrograph tape-recorder and a Leak speaker. The speaker was placed in the right-hand corner of the room in which the subject sat, facing inwards towards the room's centre and in front of the subject at a distance of approximately 1·44 m. from him. An attempt was made to make the musical selections 'stimulating' in accordance with criteria suggested by Wokoun (1969). All the selections were instrumental ones, lasting an average of 2 min., and ranged from a solo guitar performance (Laurindo Almeida) to a large modern orchestra (Don Ellis). A number of measures relating to subjective feelings during the performance of the task were taken, but since these do not directly concern the effects of music upon detection efficiency, the findings will be reported elsewhere. Virtually all subjects found the music stimulating, but not all of them liked the musical selections. In the 'noise' conditions white noise of approximately the same intensity (75 dbA) was relayed to the subject in the same way as that described above. This was the comparison condition used by McGrath (1963).

Subjects were informed that the experiment was concerned with the effects of music upon working efficiency at a monotonous task. They were given a 5 min. sample of the task, in which six signals were presented at the same average inter-signal interval as that used in the main experiment, and were instructed to respond by pressing the appropriate key whenever a signal or standard light was presented. The signal was described to them and subjects began the main experiment once they had detected all the signals in the training session. No subject needed more than one session to achieve this. Subjects' watches were removed and they were given no indication of how long the task was to last.

Table 1. *The mean percentage of signals correctly detected in each of the* 10 *min. periods of the task in the four conditions of the experiment*

	10 min. periods of task			
Condition ...	1	2	3	4
Noise easy	87	88	92	88
Music easy	86	87	89	87
Noise difficult	82	76	73	80
Music difficult	94	94	92	84

Results

Correct detections, commission errors (that is, responding to a standard light as if it were a signal) and detection latencies were the principal measures taken. Correct detection and detection latency scores were each analysed in a 2 (easy/difficult) × 2 (music/noise) × 4 (time periods) analysis of variance. Commission error scores were analysed non-parametrically, since their distributions tended to be skewed. In addition, d' measures were derived and were also subjected to a 2 × 2 × 4 analysis of variance. The results for these four response measures are treated separately below.

Correct detections. The mean percentage of correct detections, by condition and by time, are shown in Table 1. In the easy conditions, more signals were correctly detected in noise (89 per cent) than in music (87 per cent), while in the difficult conditions more were detected in music (91 per cent) than in noise (73 per cent). However, the only main effect or interaction which proved to be significant was the music × difficulty × time periods interaction ($F = 3{\cdot}96$; d.f. = 3, 108; $P < 0{\cdot}05$). This significant interaction can perhaps best be interpreted as indicating that when the task is easy, music has virtually no effect on the detection rate, and the detection rate is not affected by time at work, performance being maintained at much the same level throughout. However, when the task is difficult, music improves the detection rate substantially, although time at work also influences the percentage of signals correctly detected. In both music and noise conditions there is a tendency for performance at the difficult task to decline with time, with the exception of the fourth time period in the noise condition, where performance improves.

Detection latencies. Mean detection latency scores for the four conditions of the experiment over the four 10 min. periods of the task are shown in Fig. 1, and the results of the 2 × 2 × 4 analysis of variance are shown in Table 2. As can be seen from Fig. 1 and Table 2, there is a significant main effect of difficulty ($F = 4{\cdot}71$; d.f. = 1, 36; $P < 0{\cdot}05$) with the difficult conditions producing longer detection latencies, and a significant interaction of music × difficulty × time periods ($F = 9{\cdot}07$; d.f. = 3, 108; $P < 0{\cdot}001$). This significant interaction would seem to indicate that under difficult conditions in noise detection latencies significantly increase as a function of time on task, while in no other condition does time at work exert a significant influence on detection latencies. As with correct detections, when the task is easy the addition of music makes little difference to performance, but when the task is difficult, music improves performance quite considerably. This effect is much more clear-cut in the case of detection latencies, however.

Commission errors. The mean number of commission errors made in the easy

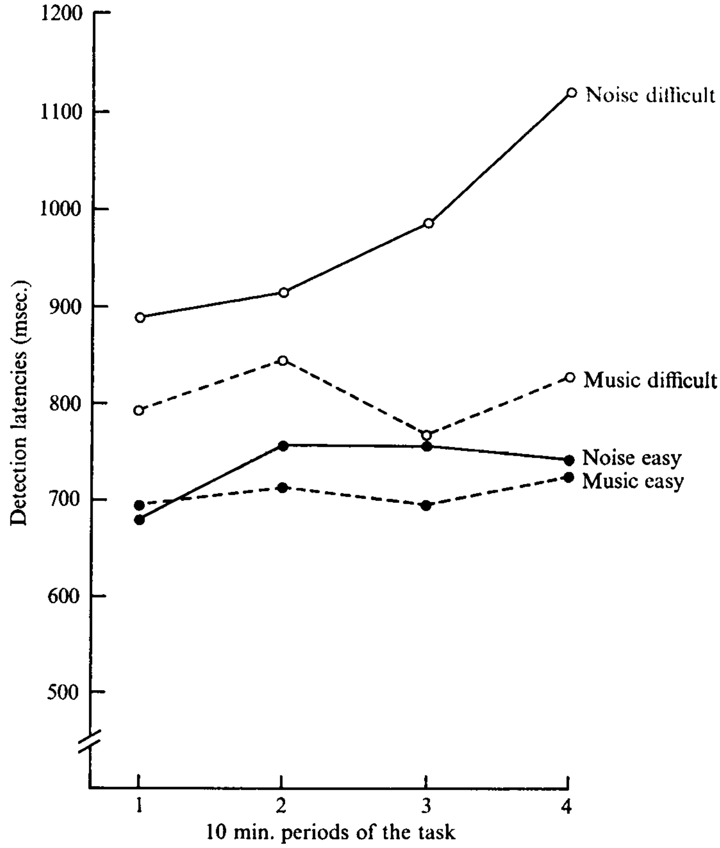

Fig. 1. Mean detection latencies in each of the 10 min. periods of the task in the four conditions of the experiment.

Table 2. *Summary of 2 × 2 × 4 analysis of variance for detection latencies*

Source of variation	S.S.	D.F.	M.S.	*F*	*P*
Difficulty	1,210,898	1	1,210,898	4·71	0·05
Music	366,911	1	366,911	1·45	n.s.
Difficulty × music	307,884	1	307,884	1·21	n.s.
Error (*a*)	9,113,225	36	253,145		
Trials	152,007	3	50,002	0·62	n.s.
Difficulty × trials	313,648	3	104,549	1·31	n.s.
Music × trials	102,850	3	34,250	0·43	n.s.
Music × difficulty × trials	2,161,578	3	720,526	9·07	0·001
Error (*b*)	8,573,809	108	79,387		

conditions was greater for music (12·40) than for noise (2·90), although the difference did not prove to be significant (Mann–Whitney U test, $P > 0{\cdot}10$). Inspection of the raw data revealed that of the 124 commission errors made in the ME condition, 109 were contributed by one subject. Commission errors were more evenly distributed in the difficult condition and here, using a Mann–Whitney U one-tailed test, significantly more ($P < 0{\cdot}05$) commission errors were made in music (98·3) than in noise (38·0). In the difficult conditions, the commission error rate declined significantly

Table 3. *Mean d′ values for each of the* 10 *min. periods of the task in the four conditions of the experiment*

	10 min. periods of task			
Condition ...	1	2	3	4
Noise easy	4·673	4·516	5·648	4·852
Music easy	4·455	4·596	4·212	4·725
Noise difficult	3·038	3·005	2·853	3·632
Music difficult	2·779	2·940	3·025	2·478

with time at work (Wilcoxon T tests, $P < 0{\cdot}05$ in each case) while no such decline was apparent in the easy conditions, where the error rate remained fairly stable throughout.

d′ measures. d′ scores, by condition and by time, are shown in Table 3. Values of d' for each subject for each 10 min. subperiod of the task were calculated using Freeman's tables (Freeman, 1964). In cases where a subject made no false positive responses during a subperiod, the lowest probability of making a false positive given by Freeman, 0·0001, was used to arrive at a value of d'. This procedure is a very approximate one, but it does allow comparisons to be made between the difficult and easy conditions of the present experiment. However, since such a procedure yields very large values of β, which are markedly different from those obtained in psychophysical experiments (Green & Swets, 1966), and in any case the assumption of the operating characteristic would be arbitrary, analyses of β were not proceeded with. The results of the $2 \times 2 \times 4$ analysis of variance indicated that there was a significant effect of difficulty ($F = 14{\cdot}93$; d.f. $= 1, 36$; $P < 0{\cdot}001$) but no other main effect or interaction proved to be significant. As might be expected, values of d' are significantly lower when the task is difficult than when it is easy.

Discussion

The results of the present study suggest, firstly, that for this task, the measure most sensitive to the effects of an environmental variable is detection latency. This measure is also the only one to show a typical 'vigilance decrement' and then only in noise in the difficult version of the task. Music abolishes this decrement, however – a result that is in accord with the findings of Wokoun (1963, 1969), although he did not vary task difficulty. Broadly similar results, although the effect is less marked, are found for correct detections. Neither the finding concerning detection latencies nor that concerning correct detections are in accord with McGrath's (1963) finding that in a more difficult version of a sensory vigilance task music no longer improves, but tends to impair, monitoring performance as measured in that experiment by detection rate. However, as noted previously, McGrath made his task more difficult by increasing stimulus presentation rate and reducing stimulus duration, rather than by reducing the intensity difference between signal and non-signal.

When experiments which have measured detection latencies to signals of different intensities are examined, it appears that low intensity signals (at or near threshold) are associated with progressively longer detection latencies as a function of time on task. Lisper *et al.* (1972), for example, using a 2 hr. auditory monitoring task, found that despite an extremely high signal frequency of 240/hr., detection latencies to threshold signals showed a sharp increase over time, while signals of higher intensities showed a much slower rate of increase. The most intense signals used (88 db) were associated with the shortest reaction times throughout and the threshold signals the longest. In a similar experiment, also using an auditory monitoring task, Loeb & Schmidt (1963) found that the progressive increase in reaction times to low intensity (10 db) signals could be prevented by administering true knowledge of results by means of signal lights which informed the subject whether the response he had just made was faster or slower than his previous response. Pseudo feedback, bearing no relation to actual speed of response, was also found to be helpful in preventing an increase in latency, although its effect was not as marked as was that of genuine feedback.

It has been suggested that knowledge of results provides information about the range of inter-signal intervals used in a particular vigilance task and hence enables the observer to build up expectancies about when signals will occur (Baker, 1963). However, since false feedback often exerts a significant effect upon indices of vigilance performance, as in Loeb and Schmidt's experiment, it appears that the administration of knowledge of results, whether true or false, tends to produce an enhancement of motivation. Thus the effects of knowledge of results upon performance in vigilance and related tasks are likely to be due in part to incentive effects which result in increases in arousal (Poulton, 1966). Knowledge of results cancels to some extent the effects of sleep deprivation but augments the effects of loud noise (100 db) upon performance at the five-choice serial reaction task (Wilkinson, 1961, 1963).

The best interpretation of the present results is probably also in terms of arousal. The effect of music upon detection latencies in the more difficult version of the task is the same as the effects of true and, to a lesser extent, false knowledge of results in Loeb & Schmidt's (1963) experiment. Since it appears that a possible interpretation of their results can be made in terms of arousal, this explanation may also be advanced for the effects of music in the present study. The beneficial effects of music only appear with the difficult version of the task, which also employed a less intense signal. It thus seems that either a more intense signal, or music, can cancel out the decrement in detection latencies.

Finally, the fact that there is a main effect of difficulty upon detection latencies, with detection latencies to more difficult and/or less intense signals being greater, is in accord with previous work. It is reasonable to suppose that in the difficult version of the task it was the brightness discrimination between the signal and non-signal that proved difficult, as implied by the main effect of difficulty upon d' values. This also results in more commission errors being made, and music appears to increase this tendency, although this finding contrasts with that of Davies *et al.* (1969). However, their task was a cognitive one, and the effect on commission errors only appeared in more extraverted subjects. The present task is a sensory one and the results were not analysed in terms of temperament. Differences in commission error rates between different types of vigilance task would appear to warrant further investigation.

This research was supported by grant HR 833/1 from the Social Science Research Council. A draft of this paper was completed while the senior author was visiting the Centre de Psychologie Expérimentale et Comparée at the University of Louvain. The authors wish to thank David Mullany and Christopher Peace for valuable discussion.

References

BAKER, C. H. (1963). Further towards a theory of vigilance. In D. N. Buckner & J. J. McGrath (eds.), *Vigilance: a Symposium*. New York: McGraw-Hill.

BUCK, L. (1966). Reaction time as a measure of perceptual vigilance. *Psychol. Bull.* **65**, 291–308.

DAVIES, D. R., HOCKEY, G. R. J. & TAYLOR, A. (1969). Varied auditory stimulation, temperament differences and vigilance performance. *Br. J. Psychol.* **60**, 453–457.

DAVIES, D. R. & TUNE, G. S. (1970). *Human Vigilance Performance*. London: Staples.

FREEMAN, P. R. (1964). Tables of d' and β. (Medical Research Council, Applied Psychology Research Unit Report, no. 529/64.)

GREEN, D. M. & SWETS, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

LISPER, H.-O., KJELLBERG, A. & MELIN, L. (1972). Effects of signal intensity on increase of reaction time on an auditory monitoring task. *Percept. mot. Skills* **34**, 439–444.

LOEB, M. & SCHMIDT, E. A. (1963). A comparison of the effects of different kinds of information in maintaining efficiency on an auditory monitoring task. *Ergonomics* **6**, 75–82.

MCGRATH, J. J. (1963). Irrelevant stimulation and vigilance performance. In D. N. Buckner & J. J. McGrath (eds.), *Vigilance: a Symposium*. New York: McGraw-Hill.

POULTON, E. C. (1966). Engineering psychology. *Ann. Rev. Psychol.* **17**, 177–200.

TARRIÈRE, C. & WISNER, A. (1962). Effects of the administration of meaningful and meaningless noise during a vigilance task. *Travail humain* **25**, 1–28.

WARE, J. R., KOWAL, B. & BAKER, R. A. (1964). The role of experimenter attitude and contingent reinforcement in a display. *Hum. Factors* **6**, 111–115.

WHITTENBURG, J. A., ROSS, S. & ANDREWS, T. G. (1956). Sustained perceptual efficiency as measured by the Mackworth 'clock' test. *Percept. mot. Skills* **6**, 109–116.

WILKINSON, R. T. (1961). Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. *J. exp. Psychol.* **62**, 263–271.

WILKINSON, R. T. (1963). Interaction of noise with knowledge of results and sleep deprivation. *J. exp. Psychol.* **66**, 332–337.

WOKOUN, W. (1963). Vigilance with background music. (Aberdeen Proving Ground, Maryland. Human Engineering Lab. Rep. No. TM-16-63.)

WOKOUN, W. (1969). Music for working. *Sci. J.* **5** (11), 33–38.

(*Manuscript received* 1 *July* 1972)