

AD 664749

22

U. S. ARMY

Technical Memorandum 1-68

EFFECTS OF MUSIC ON WORK PERFORMANCE

William Wokoun

DDC  
RECORDED  
FEB 9 1968  
RECEIVED  
D

January 1968

AMCMS Code 5026.11.81900

HUMAN ENGINEERING LABORATORIES



ABERDEEN PROVING GROUND,  
MARYLAND

This document has been approved for public  
release and sale; its distribution is unlimited.

Reproduced by the  
CLEARINGHOUSE  
for Federal Scientific & Technical  
Information Springfield Va. 22151

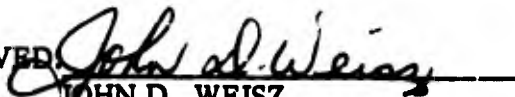
46

EFFECTS OF MUSIC ON WORK PERFORMANCE

William Wokoun

January 1968

APPROVED



JOHN D. WEISZ

Technical Director

Human Engineering Laboratories

U. S. ARMY HUMAN ENGINEERING LABORATORIES  
Aberdeen Proving Ground, Maryland

This document has been approved for public  
release and sale; its distribution is unlimited.

## ABSTRACT

Sixty-three subjects worked at a vigilance task for an hour while listening to one of three musical programs. While all three programs included the same 23 selections, the Ascending Program grew steadily more lively, and the Descending Program grew steadily less lively, while the Increasingly Variable Program progressively increased the contrast between adjacent selections. Results showed that changing the sequence of the 23 selections profoundly affected reaction times and variabilities, as well as individual consistency. Subjects performed the vigilance task better with the Ascending Program than with the other two programs.

## CONTENTS

ABSTRACT . . . . .	iii
INTRODUCTION . . . . .	1
METHOD	
Subjects . . . . .	5
Experimental Rooms . . . . .	5
Music . . . . .	5
Vigilance Task . . . . .	11
Control and Data-Recording Equipment . . . . .	12
Data Conversion . . . . .	16
Procedure . . . . .	16
RESULTS AND DISCUSSION	
Response Measures and Criteria . . . . .	20
Mean Response Times . . . . .	20
Variability of Response Times . . . . .	25
Longest Response Times . . . . .	28
More-Stimulating vs. Less-Stimulating Selections . . . . .	29
Consistency (Reliability) of Individual Performance . . . . .	30
Attitude-Scale Scores . . . . .	32
Conclusions . . . . .	33
SUMMARY . . . . .	34
REFERENCES . . . . .	35

**FIGURES**

1. Secretary Demonstrates How Subjects Watched Vigilance Box for Stimuli . . . . . viii

2. Vigilance Box . . . . . 10

3. Stimulus-Control Console . . . . . 14

4. Data-Storage Console . . . . . 15

5. Mean Reaction Times, by Quarter Hours, for Ascending and Descending Programs . . . . . 22

6. Mean Reaction Times, by Quarter Hours, for Increasingly Variable Program, as Compared to Ascending and Descending Programs . . . . . 24

**TABLES**

1. Ascending Program . . . . . 6

2. Descending Program . . . . . 8

3. Increasingly Variable Program . . . . . 9

4. Attitude Toward Music During the Experiment . . . . . 18

5. O'Neill's Classification of the 23 Selections . . . . . 23

6. Means and Variances for the Three Programs, by Quarter Hours . . . . . 25

7. F Ratios Comparing Variabilities Between Quarters of Programs . . . . . 26

8. F Ratios Comparing Variabilities Between Programs, by Quarters . . . . . 27

9. Longest Responses: Means, Variances, and Distribution by  
Quarters . . . . . 28

10. Correlation Between Response Times, by Half Hours . . . . . 30

11. Correlation Between Response Times, by Quarter Hours . . . . . 32



Fig. 1. SECRETARY DEMONSTRATES HOW SUBJECTS WATCHED VIGILANCE BOX FOR STIMULI

# EFFECTS OF MUSIC ON WORK PERFORMANCE

## INTRODUCTION

A preceding study (19) demonstrated that subjects worked significantly more effectively at a vigilance task while listening to functional music -- music selected and programmed to help them maintain their alertness -- than while hearing comparably loud noise. But while it is clear that subjects performed better, there remains a question of generality: would they have done just as well with any other musical program? Just what was it that improved performance -- simply the fact that there was music, or a particular kind of music, or a certain sequence of selections, or what?

There is already considerable research comparing performance at assorted tasks, with music and without music. It has created a veritable jungle of principles and conclusions which conflict wildly, probably because so many different kinds of music have been used. As an aggregate, these studies argue rather conclusively that the effects music will have on behavior depend on the kind of music used. Thus the first possibility -- that any music will automatically improve performance -- must be discarded.

Besides merely comparing music with silence or noise, some research has also attempted to relate type of music to performance changes.

One group of experimenters has classified music by the behavioral effects they think it will have. For instance, Skelly and Haslerud (15) report that apathetic schizophrenics are more active while listening to "livelier" music. Shatin, Lussier, and Kotter (14) found that "stimulating" music calmed fearful schizophrenics more than quieter music or silence. With ratings of pastel drawings as a criterion, Offer and Stine (11) noted that both depressed patients and control subjects responded differently to stimulating music, relaxing music, and no music. Weidenfeller and Zimny (18) extended this finding by measuring galvanic skin responses (GSRs) of depressed and schizophrenic patients to "exciting" and "calming" music. Music affected GSRs in both groups, although the depressive patients showed greater response to "exciting" music.

To find out more about normal subjects' behavior, Zimny and Weidenfeller (21) then checked how college students' GSRs and heart rates varied while they listened to three selections that differed in "exciting" quality. The more "exciting" selections did seem to affect GSRs, but not heart rates. This result parallels an earlier study by Ellis and Brighthouse (5), which had demonstrated that both "Hungarian Rhapsody" and "Blue Interval" increased respiration rates, although they had no significant effect on heart rates. These studies prove that selections representing the extremes of musical dimensions do affect at least some kinds of behavior.

Rather than selecting music for its hypothesized effects, other experimenters have classified their music by schools and periods. For example, Freeburne and Fleischer (6) exposed their subjects to four types of music -- classical, semi-classical, popular, and jazz -- as well as a control condition of silence. Subjects who heard jazz read significantly faster than the other groups, although their comprehension scores were no better. Trazel and Wrede (16) measured GSRs while students listened to three types of music: symphony, operetta, and jazz. Once again, it was jazz music that affected behavior most, even though the subjects said they preferred symphonic music.

But differences in music do not invariably produce corresponding changes in behavior. Van Krevelen (17) reported that college students answered a frustration questionnaire the same way with melodic symphonic music as with dissonant symphonic music -- but differently when there was no music. Nelson and Finch (10) had subjects ride a bicycle ergometer for one minute under each of four conditions: fast march music, slow music, white noise with mambo background, and no sound. There were no significant differences among the four conditions.

These inconsistencies suggest that music is too complex, or too variable, or simply embodies too many attributes, to be classified meaningfully by just its period or musical form.

Henkin (7) attempted to pinpoint the significant variables of music through factor analysis. Although he selected his music by sampling musical periods -- baroque, pre-classic, classic, romantic, and modern -- he treated the esthetic judgments mathematically to extract their independent, underlying variables. There were two clear-cut factors -- melody and rhythm -- and the suggestion that orchestral color might be a third factor. Later Henkin (8) detected two additional factors: a second melodic factor, and one which he did not name.

Then Henkin (9) related his factor analytic results to GSR patterns. He found that GSRs are influenced by both the melodic and the rhythmic factors. Interestingly, GSRs did not seem to depend on musical style, dynamics, orchestration, timbre, or other compositional techniques.

Similarly, Cattell and Anderson (3) factor-analyzed judgments of 120 musical excerpts, and their work was extended by Cattell and Saunders (4) to isolate 12 factors. The experimenters interpret these variables as "personality and temperament factors." As in the Henkin studies, these factors seem independent of musical periods and schools.

All of these studies underline the fact that the nature of the music itself is a critical variable. Yet a vital link is missing: virtually all of the experiments in the literature have used recordings of music intended for entertainment, rather than music tailored specifically for functional use. Since functional music is selected, arranged, and recorded to meet entirely different criteria, it would reasonably be expected to have different effects on behavior. Thus the studies in the literature indicate how isolated specific selections affect performance. Indirectly, they may well indicate which experimenters could develop effective musical programs intuitively. But unfortunately, they do not show how the specially arranged selections usually used in work environments affect performance, nor do they indicate what the sequence of selections should be like.

Complementing these studies of individual selections, other experimenters have attempted to develop theoretical models of how music affects performance. The earliest attempt to modify listeners' behavior by manipulating variables within the musical program seems to have been Wyatt and Langdon's (20) work shortly before World War II in Great Britain. Their principles have been extended by Cardinell (1), Cardinell and Burris-Meyer (2), and most explicitly by O'Neill (13).

In brief, O'Neill maintains that music has a stimulation value which is not unlike Hull's concept of stimulus-intensity dynamism (V). According to O'Neill, any performance of a musical selection has its inherent stimulation value, or net effect on the worker's alertness and performance; for the moment, the term "stimulation value" is used in this sense without attempting to specify underlying physiological dimensions prematurely. These stimulation values depend on a number of musical variables, such as tempo, rhythm, instrumentation, and orchestral size, as well as others that are difficult to quantify. Some of the variables which O'Neill names -- such as rhythm and orchestral color -- parallel the ones which emerged from Henkin's factor analysis.

O'Neill maintains that these stimulation values reside not so much in a melody itself as in the way it is arranged and recorded; the same melody can be made more or less stimulating by manipulating its tempo or any of various other musical variables. An arrangement's stimulation value also depends on contrast, i.e., on the stimulation values of the selections which precede and follow it.

By this reasoning, the best musical program would be one that counterpoises stimulation values and work decrements, using more stimulating music to compensate for decreases in work output. While it is possible that work output might increase during a shift, it is much more typical for work-output curves to show progressive degradation, at least up to a point, with a tendency to recover near the end of the shift. Thus O'Neill concludes that ascending programs -- those in which the selections become gradually but continuously more stimulating -- are most effective in industrial work situations. By extrapolation, there should be a similar relationship in vigilance tasks, although we know considerably less about the shape of the "work-output" curve for vigilance.

Most fundamentally, then, O'Neill states that the music which his criteria rate more stimulating will have a different effect on behavior than music rated less stimulating. This hypothesis can be tested experimentally in a work situation much like the one used previously. A group of recorded functional-music selections can be ranked on stimulation values, then permuted in several ways to give musical programs which administer exactly the same selections to every subject, but in different orders -- and, consequently, with maximum stimulation at different times.

If O'Neill's hypothesis is correct, these permuted programs should produce different reaction-time curves. We need not predict which program will give "best" performance, and it seems rather unlikely that an exploratory study would stumble onto a "best possible" program. The basic points this experiment aims to examine are (a) whether the ranked stimulation values of musical selections do affect performance, and (b) whether their sequence (order) is important. If so, these facts would support O'Neill's hypothesis. But if they do not, we must conclude that the stimulation values, and the order in which they are administered, do not affect work performance.

## METHOD

### Subjects

The subjects were 63 enlisted personnel from the 14th Enlisted Training Company and the Officers Candidate School at Aberdeen Proving Ground, temporarily assigned to the Human Engineering Laboratories as subjects.

### Experimental Rooms

Each subject was tested while seated alone in an Industrial Acoustics Co. model 402-A audiometric booth. These booths effectively isolated the subjects from most of the extraneous noises they would otherwise have heard. Inside, the booths were approximately six feet square and 6 1/2 feet high. A vigilance stimulus box was on the table in front of the subject. To his left, on the floor, there was a seven-inch loudspeaker in a closed baffle.

### Music

The basic pool of musical selections was the second hour of a program which Muzak prepared for a vigilance experiment several years ago (Muzak serial no. Z-35069). These selections were not entertainment music, but music arranged and recorded specifically for functional use. None of these 23 selections was used in the previous experiment. The 23 titles were submitted to O'Neill, who arranged them in rank order from least stimulating to most stimulating. It would, of course, be highly desirable to measure stimulation value, rather than merely ranking it. Unfortunately, there seems to be no way of expressing over-all stimulation as a single number, at least at present.

Note, however, that any inaccuracies in ranking would tend to reduce the differences among the programs. If the ranking were purely random, average stimulation value would be roughly the same in each quarter hour, and all programs should give the same reaction-time curves. Hence inaccurate rankings could hardly induce a spurious difference, although they might well cause negative results.

TABLE 1  
Ascending Program

Music Program		Stimuli	
Start Time	Selection	Time	Stimulus Tube
:00	But Beautiful	1:45	4
2:31	Artificial Flowers	4:22	1
5:18	Will You Still be Mine		
7:56	The Best Things in Life Are Free	8:06	3
10:29	Mexicali Rose	12:25	2
13:01	By the Bend of the River		
15:31	I Left My Heart in San Francisco	16:44	4
17:45	No Other Love		
20:27	Tumbling Tumbleweeds	21:58	2
22:38	Two Cigarettes in the Dark	25:19	2
25:41	Till Tomorrow		
28:22	Love Look Away	29:03	1
30:59	Auf Wiedersehen	31:51	3
34:08	Souvenir d'Italie	35:56	2
36:38	Beautiful People of Denver		
38:33	You and I	38:42	4
41:42	Nothing More to Look Forward to	42:33	3
44:29	On the Side of the Angels	46:00	4
46:53	Hernando's Hideaway	48:33	1
50:04	Love is Here to Stay	52:56	3
53:24	Learnin' the Blues		
56:07	Tea for Two	57:08	1
58:45	Midnight in Paris	60:01	2
61:19	(End of program)		

After the selections had been ranked, they were arranged into three programs:

1. The Ascending (A) Program began with the least stimulating selection ("But Beautiful") and grew progressively more stimulating, ending with the most stimulating selection ("Midnight in Paris"). This order is the same sequence that O'Neill prepared.

2. The Descending (D) Program used exactly the opposite order: it began with the most stimulating selection ("Midnight in Paris") and grew steadily less stimulating.

3. Since stimulation value changed gradually in the first two programs, it seemed worthwhile to compare them with a third program that would maintain roughly the same average stimulation during the hour, but do so by alternating more-stimulating and less-stimulating selections, and increasing the contrast between adjacent selections. This Increasingly Variable (V+) Program began with the selection at the middle of O'Neill's rank order ("Love Look Away"), then the one below it, then the one above it, and so on. It ended with the most extreme contrast the group of 23 selections allowed ("But Beautiful," followed by "Midnight in Paris").

The sequences of the 23 basic selections in these three programs are shown in the first two columns of Tables 1, 2, and 3.

After the three program sequences had been planned, each program was recorded at 3 3/4 inches per second on a 1163-foot Fidelipac tape cartridge. All 23 selections were played from discs, electronically compressed into a dynamic range of 20 decibels (dB) to restrict variations in loudness, and recorded on a master tape. Each selection was then located on the master tape and played individually so it could be transferred to each of the three cartridges. Once the controls on the playback and recording amplifiers had been set, they were left undisturbed until all three cartridges had been completed. Thus the signals recorded on the different cartridges were as nearly identical as possible. The same three cartridges were used throughout the experiment.

During the experiment, the subjects heard the music program over their loudspeakers. The loudness of the music at the subject's ear position ranged between 55 and 65 dB re 0.0002 microbar, as measured with a Bruel and Kjaer type 2203 Precision Sound-Level Meter.

TABLE 2

## Descending Program

Music Program		Stimuli	
Start Time	Selection	Time	Stimulus Tube
:00	Midnight in Paris	2:18	3
2:36	Tea for Two	5:03	1
5:14	Learnin' the Blues		
7:57	Love is Here to Stay	10:22	4
11:16	Hernando's Hideaway		
14:26	On the Side of the Angels	14:26	2
16:52	Nothing More to Look Forward to	17:49	1
19:36	You and I	21:34	3
22:46	Beautiful People of Denver		
24:40	Souvenir d'Italie	25:16	2
27:13	Auf Wiedersehen	28:42	4
30:22	Love Look Away	31:16	3
32:58	Till Tomorrow	34:32	4
35:39	Two Cigarettes in the Dark	38:39	2
38:43	Tumbling Tumbleweeds		
40:56	No Other Love	42:20	2
43:38	I Left My Heart in San Francisco		
45:53	By the Bend of the River	46:27	1
48:25	Mexicali Rose	50:11	4
50:58	The Best Things in Life Are Free		
53:32	Will You Still be Mine	54:46	1
56:11	Artificial Flowers	58:10	3
58:57	But Beautiful	60:48	2
61:26	(End of program)		

TABLE 3

## Increasingly Variable Program

Music Program		Stimuli	
Start Time	Selection	Time	Stimulus Tube
:00	Love Look Away		
2:38	Till Tomorrow	3:18	2
		5:10	1
5:20	Auf Wiedersehen		
8:29	Two Cigarettes in the Dark	9:30	4
11:34	Souvenir d'Italie	13:35	1
14:07	Tumbling Tumbleweeds		
16:21	Beautiful People of Denver	16:49	3
18:16	No Other Love	19:24	4
20:58	You and I	23:47	3
24:07	I Left My Heart in San Francisco		
26:22	Nothing More to Look Forward to	27:40	2
29:08	By the Bend of the River	31:22	4
31:40	On the Side of the Angels		
34:07	Mexicali Rose	35:28	1
36:39	Hernando's Hideaway	39:11	3
39:51	The Best Things in Life Are Free		
42:27	Love is Here to Stay	43:03	2
		45:32	2
45:48	Will You Still be Mine		
48:28	Learnin' the Blues	48:47	4
51:13	Artificial Flowers	51:53	1
53:59	Tea for Two	55:37	3
56:37	But Beautiful	58:45	1
59:10	Midnight in Paris		
61:43	(End of program)		

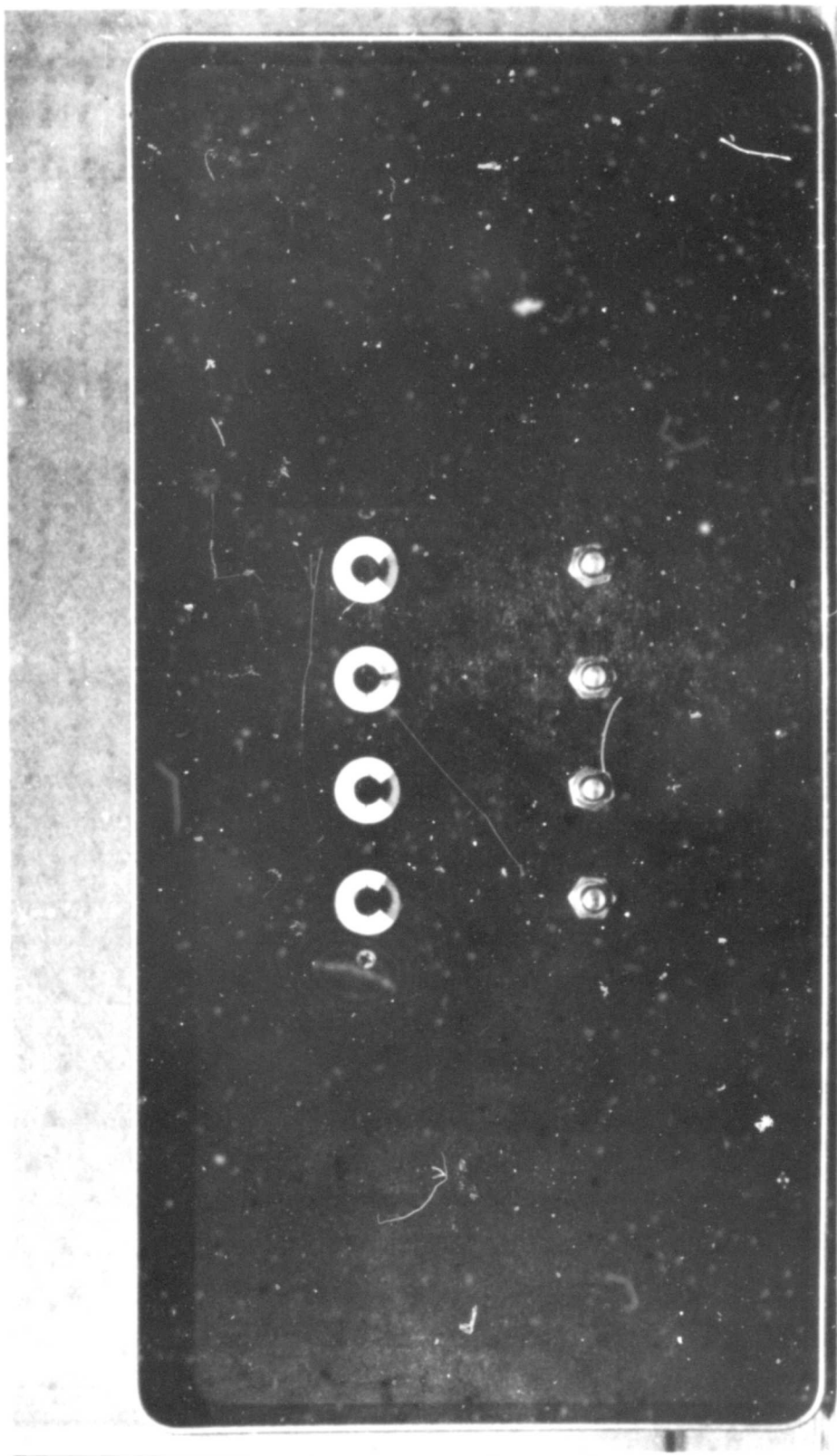


Fig. 2. VIGILANCE BOX  
(Third tube from left displays a stimulus.)

## Vigilance Task

In the previous study, subjects watched a square of translucent plastic that could be illuminated by any of three colored lights, and responded by pressing one of three telegraph keys. While this apparatus is excellent for measuring simple reaction times, the light seemed too glaringly obvious to be a suitable vigilance stimulus. Because the light was so bright, the subjects could see it even when they were not looking directly at it. Even if the light had been dimmed, the subject's task would still be rather undemanding, since every stimulus would still have to appear at the same place.

Modified vigilance boxes were designed for the present experiment to make the task more difficult: to make the stimuli subtler, and to force subjects to scan several stimuli, rather than to merely stare at one. This box (Fig. 2) used four 6E5 electronic tubes, commonly called "magic-eye" tubes, mounted 1 1/2 inches apart. The control grids of these tubes were normally grounded, giving a wide shadow at the bottom of the tube (from left to right, Tubes 1, 2, and 4 in Fig. 2). To display a stimulus, the tube's grid was biased with a 1 1/2-volt alkaline battery, and its shadow narrowed to about half its normal width (see Tube 3 in Fig. 2). While this narrowing did increase the tube's brightness slightly, the change in intensity was not a dramatic one, and subjects had to watch the tubes to detect it. And, since they were presumably within arm's-length of the box (roughly 32 inches for an average person), they could not watch all four tubes in a single fixation.

Immediately below each stimulus tube, there was a pushbutton. The subjects were instructed to press these buttons to show they had perceived a stimulus.

Both the intervals between stimuli and the order of stimuli were randomized, with certain restrictions. It was stipulated that 17 stimuli would be given during the one-hour sessions, spaced no closer together than 1 1/2 minutes, and no farther apart than 5 1/2 minutes. Subject to these limitations, the exact times were obtained from a table of random numbers, independently for each program.

The basic sequence of the four display tubes was also determined from random numbers, with the stipulations that a given stimulus must not occur more than twice in a row, and that each stimulus must occur about equally often.

The stimulus times and sequences are shown, in relation to the music programs, in the third and fourth columns of Tables 1, 2, and 3.

## Control and Data-Recording Equipment

We have already seen that the music programs were recorded on magnetic-tape cartridges. To guarantee that every stimulus was given at exactly the same point during the session, the experimental design required that the program cartridges also control the stimuli automatically.

However, translating this seemingly simple requirement into apparatus proved to be a difficult and time-consuming feat; it is briefly summarized here for the benefit of other experimenters who may wish to replicate this experiment. The simplest approach is recording both music and tones on a single track of the tape. By choosing a suitably high frequency -- that is, one near the upper limit of audibility, where both ears and loudspeakers are less sensitive -- it should be possible to insert a control tone which can be detected neatly with tuned circuits, yet one which can conveniently be deleted from the music program by curtailing the playback system's high-frequency response. The difficulty is that tones approximating 15,000 cycles per second (cps) are very difficult to record and play back reliably. Tests with several tape recorders showed that most of them would not reproduce this tone at a reasonable level, much less a consistent one. Furthermore, since response at the extreme high frequencies depends heavily on perfect (and perfectly clean) playback heads, absolutely smooth tape, and continuous intimate contact between tape and head, there is a strong possibility that wear will make such systems unreliable.

Lower-frequency tones -- for example, 10,000 cps -- are much easier to record and reproduce reliably. The problem then becomes one of removing them from the music program so they will not give the subjects extraneous cues. Unfortunately, all filters have more or less gradual slopes of attenuation, rather than sharp, cliff-like cut-offs, so it is necessary to begin the attenuation slope at relatively low frequencies. Thus the quality of the music suffers noticeably, yet the tones remain clearly audible even though they may be 30 dB less intense than the music. By this time it seemed obvious that it was not practical to multiplex both music and tones on the same tape track.

Consequently, the monophonic head in the cartridge machine was replaced with a two-track head: one track for the music program, and the other for control tones. The two control tones were chosen for convenience: a 3700-cps tone to select which stimulus would be presented, and a 1450-cps tone to trigger the timing chain that actually presented the stimuli.

Even though the control tones were recorded on a separate tape track, cross-talk in a stereo preamplifier injected audible amounts of them into the music channel. Therefore the control-track channel of the tape playback head was connected to a completely separate preamplifier, powered by batteries and physically isolated from the one used for the music program. Several pilot subjects verified that this arrangement completely eliminated the tones from the music program, even during the silent intervals between musical selections.

The preamplified control signals were routed to a specially designed tone discriminator which amplified them further, then used highly selective telemetering filters to distinguish between them, and finally closed a corresponding relay. The stimulus-selection relay (operated by 3700 cps) cycled a stepping switch and set it to the position corresponding to the next stimulus.

The stimulus-presentation relay (operated by 1450 cps) started the timing chain for a trial by simultaneously turning on the magnetic-tape data recorder and triggering a five-second timer. Thus the data recorder had a full five seconds to stabilize speed and tape motion.

At the end of five seconds, stimuli were generated electronically by biasing the grids of the four selected display tubes with -1.5 volts. Simultaneously, a 5000-cps timing tone was gated to four tracks of the data-recorder tape, one corresponding to each booth. When a subject responded by pressing the correct pushbutton, he turned off the tone which was being recorded on his data track, and removed the negative bias from his tube so its shadow widened back to its usual size. If a subject did not respond to a stimulus after it had been presented for approximately 10.5 seconds, another timer automatically turned the stimulus off and stopped the data recorder. Thus the complete stimulus-presentation sequence took slightly more than 15 seconds.

During trials, the experimenter's console allowed him to monitor each subject's performance with pilot lights corresponding to each display tube, which came on when a stimulus was given and stayed on until the subject responded or the stimulus was terminated.

The control equipment is shown in Figure 3, and the data recorder in Figure 4.

Electromagnetic components -- such as relays, stepping switch, motors, and timers -- generated a variety of transient clicks which, if uncontrolled, could have given the subjects a rich assortment of extraneous cues. The sources of these transients were tracked down individually and neutralized with a battery of radio-interference filters, diodes, and capacitors. All transients were eliminated from the final apparatus used in the experiment, except for a single click at the end of the stimulus-presentation sequence, when it was too late to affect the subject's performance on the trial.

The timing tone was a nominal 5000-cycle sine wave generated by an Eico model 377 Audio Generator. To compensate for frequency drift, the timing frequency was measured during each trial with a Hewlett-Packard model 524B Electronic Counter.



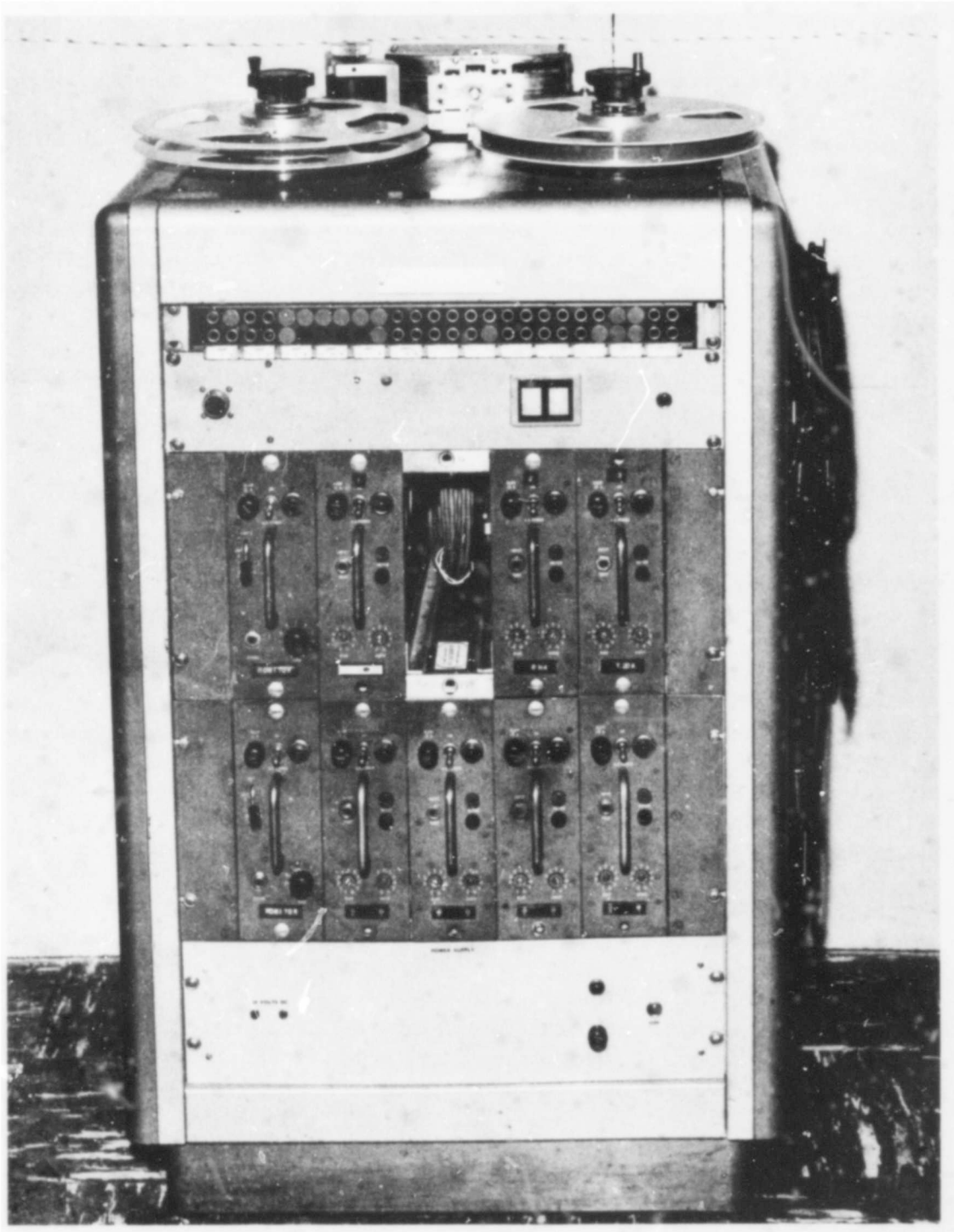


Fig. 4. DATA-STORAGE CONSOLE

Magnetic-tape data recorder stores gated timing-tone bursts that indicate subjects' response times. Patching panel at top routes tone bursts to tape heads; below it are recording and playback amplifiers, and power supply.

## Data Conversion

The product of an hour's experimental session was a half-inch magnetic tape with one track for each booth, containing both calibration tones and tone bursts whose durations measured the subjects' response times. The first step in converting tone bursts to reaction times was playing each tone burst into the Hewlett-Packard 524B Electronic Counter to get the number of timing cycles for a trial. These tone bursts were first played forward at 30 inches per second, as recorded, then spooled backward over the tape heads by hand; virtually always the two counts agreed within  $\pm 2$  cycles, although an occasional burst might require a third playing to resolve a discrepancy arising from poor tape-head contact during manual operation. A subject's cycle count for a trial was then divided by the timing frequency during that trial and rounded off, giving his response time to the nearest millisecond.

## Procedure

Before each experimental session, and without knowing which subjects the subject coordinator would supply for it, the experimenter selected the music program that would be used. The three programs were used in rotation. Since the selection of subjects was entirely independent of the experimenter, the subjects may be considered randomly assigned to the three experimental conditions (music programs).

The subjects were tested in groups of four when possible -- one subject in each of the four available booths -- although four subjects were not always available at the same time. All of the later experimental sessions were scheduled to begin at one o'clock, immediately after lunch, to minimize any effects that differing times of day might have on a subject's alertness. Early in the experiment, three groups of subjects had been tested at 2:30, one group with each of the programs. These subjects may be included in this experiment without introducing biases; since all conditions were sampled at the later time, time of day could not favor any musical program over any other.

The subject coordinator delivered the subjects to the experimental suite and introduced them to the experimenter, who asked them to select a booth and relax while waiting for instructions over the loudspeaker.

To assure that every subject received exactly the same instructions, they were recorded on a separate tape cartridge, complete with control tones that selected and presented stimuli to illustrate the directions. After each subject was seated in his booth at the beginning of the experimental session, he heard these instructions over his loudspeaker:

Modern missile systems are becoming more and more automated. Once a target has been detected, the system practically runs by itself -- it can track the target automatically, compute its course, and fire at it until it has been destroyed. But soldiers still have to keep watch on radar scopes to detect targets in the first place, and to decide which targets to fire at. This experiment will measure how well you can detect target signals.

The box in front of you has four electronic tubes in it, representing the four sectors a target can approach from. Notice that each tube has a wide shadow at the bottom -- this means there aren't any targets in that sector.

Now if a target does appear, one of the shadows will narrow to about half the size it is now. Watch. (Stimulus appears.) When a shadow does narrow, you should push the button under that tube. Push it now. (Pause)

This is how you tell the computer to track in that sector. Notice that, when you push the button, the shadow widens back to its normal size as the computer takes over.

When you see a target, push the button firmly until you hear it click, and then hold it down for a moment until you see the shadow go back to normal. Let's try it again for practice, to see how fast you can react. (Second stimulus appears.)

The important thing is pushing the button fast. Try to push the button just as quickly as you can when you see a target, because you're being scored on how fast you respond. But be careful not to press the buttons unless you do see a target. If you press a button when there isn't any target, that counts against you as an error.

Notice, too, that your booth is soundproofed, so you won't be distracted by hearing people talking, telephones ringing, and things like that. We have to use soundproofing to keep these noises under control. Our first subjects complained the booths bothered them because they were unnaturally quiet, so now we counteract that by playing background music.

Remember, when you see a target, press the button just as quickly as you can. If you have any questions, open your door and ask them now. If not, begin watching for targets as soon as the music starts playing.

A short time was allowed to give the subjects an opportunity to ask questions if they wanted to, but no subject requested further instructions. The music program itself began within a minute after the instructions were completed.

TABLE 4

Attitude Toward Music During the Experiment

Name \_\_\_\_\_

Different people feel different ways about the music that was played during part of the experiment. The statements below represent several points of view. For each statement, mark whether you:

- A - agree with it
- D - disagree with it
- ? - are undecided about it

- \_\_\_ The music made the task seem easier.
- \_\_\_ I'd have to work with music longer before making up my mind about it.
- \_\_\_ The music was really great.
- \_\_\_ The music made the task much easier for me by blocking out the noises from the hallway.
- \_\_\_ The music kept me from concentrating.
- \_\_\_ The music had no effect on me.
- \_\_\_ Sometimes I like to listen to music.
- \_\_\_ Hearing music really didn't bother me much.
- \_\_\_ I didn't like the tunes.
- \_\_\_ I just couldn't do my best without music.
- \_\_\_ The music wasn't too bad, but not too good, either.
- \_\_\_ Sometimes music makes my tasks easier.
- \_\_\_ Music is sometimes disconcerting.
- \_\_\_ The music helped me relax and do a good job.
- \_\_\_ I kept wishing they would turn the music off.
- \_\_\_ The music broke up the monotony and boredom.
- \_\_\_ The music is soothing to the nerves and makes you feel relaxed.
- \_\_\_ The music kept distracting me and caused me to lose concentration.

Thank you. If you would like to add any other comments, please write them on the back of this sheet.

Although the subjects were told they would be penalized for responding when there was no stimulus, there was no provision for counting such responses. Indeed, they would not even be expected in this unpaced vigilance situation. This instruction was inserted mainly to keep the subjects from holding buttons down continuously and recording impossibly fast reaction times. While these spuriously short times could easily be differentiated from valid response times, more-useful data would have been lost. By observing the subjects frequently through the booths' one-way windows, the experimenter satisfied himself that subjects virtually never pressed buttons unless there was a stimulus.

At the end of the hour, the experimenter asked each subject to complete an attitude test (Table 4) in his booth. Afterward, he illustrated how the subjects' reaction times had been measured and showed how the control equipment operated. In most cases this seemed to satisfy the subjects' curiosity.

Sometimes the experimenter asked the subjects rather vague questions such as, "What do you think the experiment was about?" No subject was ever able to verbalize the hypothesis under test. Still, even if isolated subjects had been able to guess the hypothesis, this knowledge would probably not have affected their performance appreciably; a recent study (12) indicates that subjects in vigilance experiments perform about the same, regardless of whether they think an experimental variable will improve or degrade their performance.

## RESULTS AND DISCUSSION

### Response Measures and Criteria

In testing the subjects' response times to determine whether the three musical programs had similar effects, there are at least three possible criteria. The most obvious index is mean response times; subjects may respond more promptly with one program than with another, or at different times during the hour.

However, we must not overlook the possibility that the programs might also affect variabilities; quite conceivably, a program could affect subjects differentially, facilitating performance for some while degrading it with others. Thus the variabilities would be important even if the means did not differ. It seems clear that a group whose response times cluster relatively closely about its mean, is preferable to one that ranges far above and below its mean.

Finally, the subjects' reliabilities -- the extent to which they retain their places in the rank order of response times throughout the hour -- give another sensitive index of differential effects. If a program does engender some differential effect, it should at least be a consistent one, so that personnel selection techniques could pick out the workers who would be facilitated. It would be decidedly inconvenient to work with a musical program which affected people so that the man who responded fastest during one part of the hour was the slowest at some other time.

From these considerations, it follows that an ideally effective music program would have these three effects:

1. Shorter mean response times.
2. Less variable response times (smaller standard deviation).
3. Higher individual reliabilities (larger correlations between a subject's performance during the two half-hour periods).

### Mean Response Times

As in the previous study, the number of stimuli was selected so individual response times could be pooled to get more stable estimates of the parameter, by giving accidental fluctuations an opportunity to offset each other. Statistically, this treatment has the added advantage that it reduces the skewness of individual response-time measures, yielding a more nearly normal distribution which can be analyzed with parametric methods.

Thus the first four trials for each program were averaged to give a mean response time for the first quarter hour, the second four trials for the second quarter hour, the third four trials for the third quarter hour, and the last five trials for the fourth quarter hour.

These quarter-hour means for the Ascending and Descending Programs are plotted in Figure 5. Response times for the Descending Program were consistently somewhat longer in each quarter hour than for the Ascending Program. The probability that this would occur by chance is the same as the probability of flipping four unbiased coins and getting four heads -- i.e., one chance in  $2^4$ , or one chance in 16, which corresponds to a probability point of .063. This is the lowest probability possible with four events.

The two curves show the general sort of relationship that O'Neill's theory would predict. More specifically, the Descending Program produced very slightly faster responses during the second quarter than it had in the first; but thereafter response times increased as the music grew less and less stimulating. Although this initial decrease is very slight, it may indicate that the subjects were learning to perform the relatively simple task better. Alternatively, beginning the session with fast, stimulating music may have tended to disrupt performance.

The Ascending Program shows response times increasing in the second and third quarters, then dropping in the final quarter hour. Although this curve seems to reflect the effect of the most stimulating music at the end of the A program, it also suggests that the music should have been more stimulating during the middle quarter hours.

This middle-quarter increase may well arise from the sample of music that was used. After listening to the selections many times, the experimenter concluded that their stimulation values increased very gradually during the first three quarter hours, then increased much more rapidly. Thus the sample of music must have had a preponderance of slow, relatively unstimulating selections, salted down with a few more lively ones. As a check, we may tabulate the one-word descriptions which O'Neill used to classify the 23 selections before he ranked them. While these labels are purely anecdotal descriptions, Table 5 shows that the distribution of selections did indeed emphasize less stimulating selections. As Table 5 demonstrates, all but two of the 23 selections were slow or medium. It should be noted, incidentally, that the experimenter was solely responsible for the choice of these 23 selections. They were chosen as they were -- as a one-hour block from an existing program -- in an attempt to avoid sampling biases. However, it would probably be wise to use stratified sampling in future studies, to assure that stimulation values do not concentrate in some areas of the continuum as they did here.

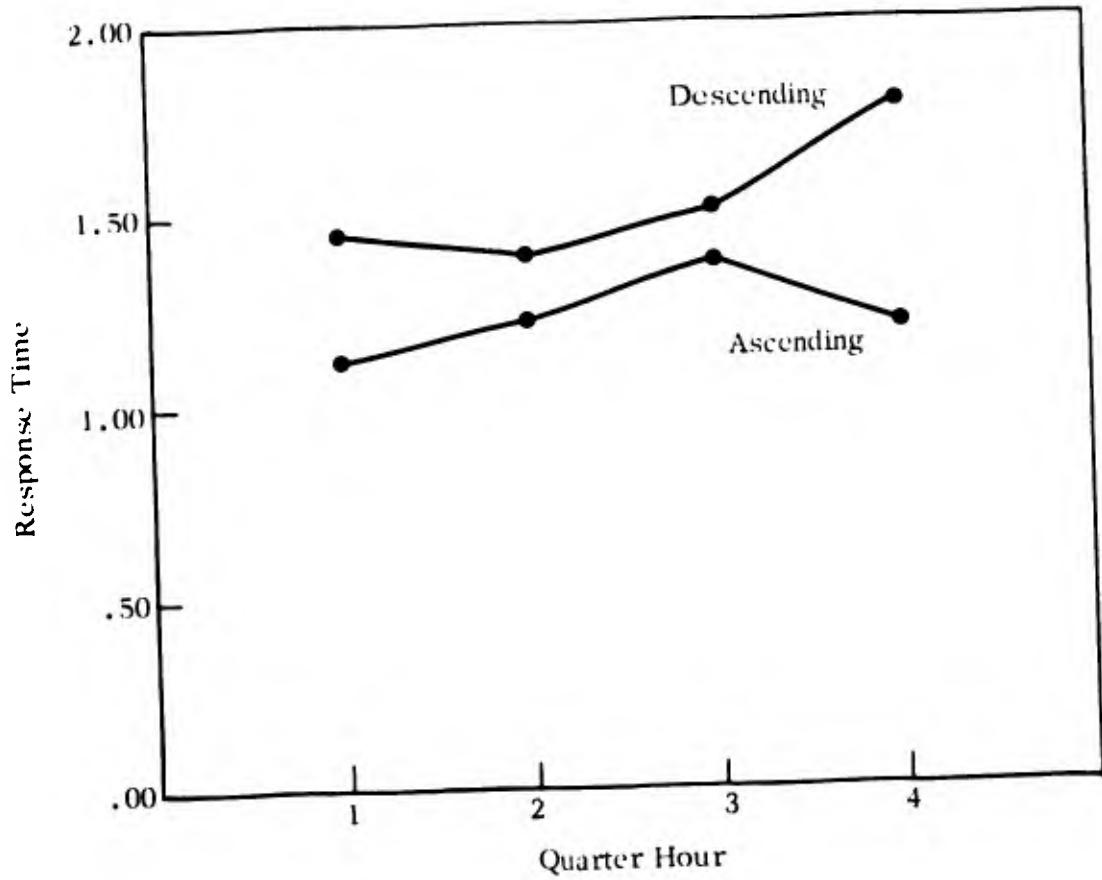


Fig. 5. MEAN REACTION TIMES, BY QUARTER HOURS, FOR ASCENDING AND DESCENDING PROGRAMS

TABLE 5

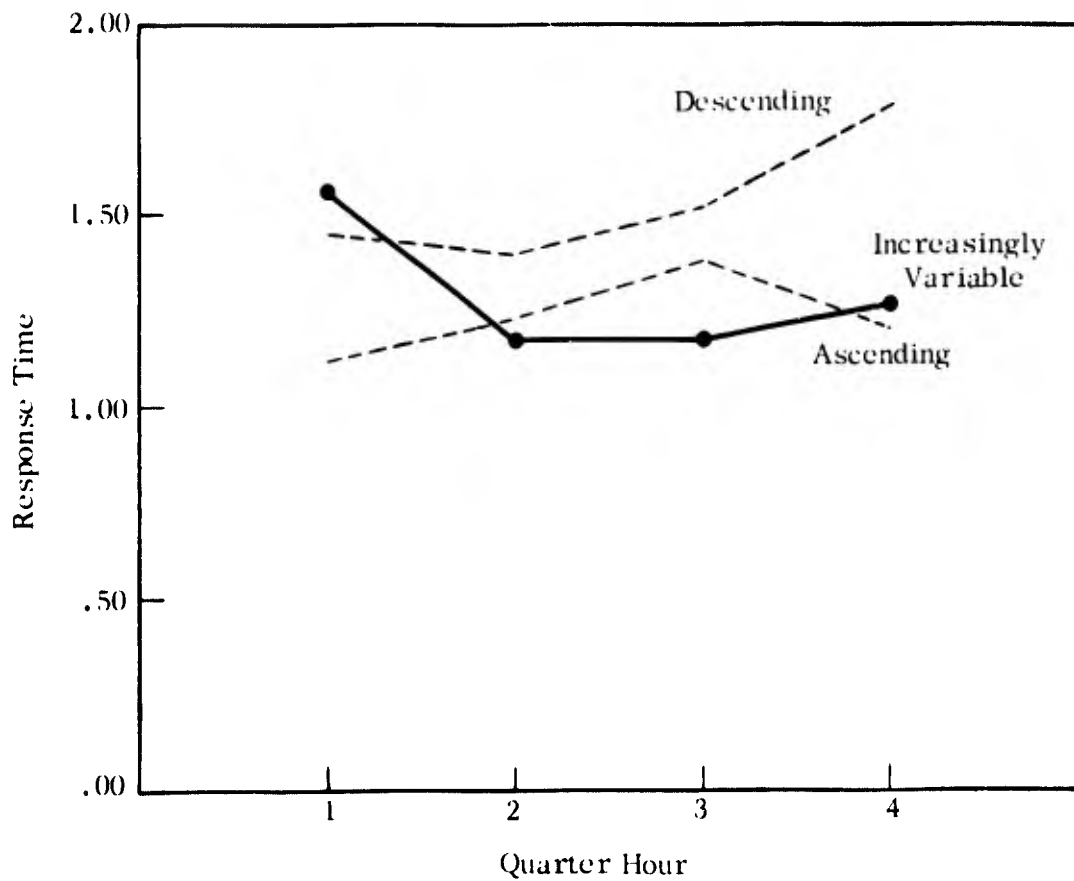
O'Neill's Classification of the 23 Selections

Classification		Number of Selections
Lively:	"Fast Waltz"	1
	"Bright"	$\frac{1}{2}$
Medium:	"Medium Waltz"	2
	"Medium +"	4
	"Medium"	$\frac{5}{11}$
Slow:	"Slow Waltz"	2
	"Slow +"	4
	"Slow"	$\frac{4}{10}$

Relating the distribution of anecdotal labels to the Ascending curve, it appears that stimulation increased too gradually to offset the vigilance decrement during the second and third quarter hours. However, the prompter responses during the last quarter hour seem to show that greater stimulation values did have their effect when stimulation increased more rapidly.

It is interesting to compare the Ascending and Descending curves by quarter hours, since the first quarter of the Ascending Program contained the same selections as the last quarter of the Descending Program, and so forth. Response times were considerably longer when these least stimulating selections came at the end of the hour. The most stimulating selections gave much more similar response times, even though the Ascending group heard them last and the Descending group heard them first.

Figure 6 relates response times for the Increasingly Variable (V+) Program to those for the other two programs. The V+ Program gave somewhat longer reaction times than the D Program during the first quarter hour, but faster ones thereafter. Compared to the A Program, the V+ Program gave faster response times during the two middle quarter hours. Over-all, the V+ and A Programs seem more similar to each other than to the D Program. This finding suggests that the subjects were stimulated not only by more lively selections, but also by greater variation or contrast between adjacent selections. It should be possible to exploit both of these principles in programming music for people working at vigilance tasks.



**Fig. 6. MEAN REACTION TIMES, BY QUARTER HOURS,  
FOR INCREASINGLY VARIABLE PROGRAM,  
AS COMPARED TO ASCENDING AND DESCENDING PROGRAMS**

## Variability of Response Times

Testing the significance of these differences requires the means and variabilities given in Table 6. The most striking -- and unexpected -- feature of Table 6 is the large range of variabilities, not only from program to program, but within each of the programs. These differences are, in fact, so large that they constitute one of the programs' major effects on performance.

TABLE 6

Means and Variances for the Three Programs, by Quarter Hours

Program	Quarter Hour	Mean (Seconds)	Variance (Seconds)
Ascending	1	1.121	.250
	2	1.238	.732
	3	1.389	.396
	4	1.200	.156
Descending	1	1.458	.856
	2	1.400	.549
	3	1.520	1.302
	4	1.744	2.259
Increasingly Variable	1	1.571	.552
	2	1.177	.316
	3	1.173	.119
	4	1.254	.309

Table 7 gives F ratios comparing the variabilities within each program, by quarters.

TABLE 7  
F Ratios Comparing Variabilities Between Quarters of Programs

Quarter Hours Compared	Programs		
	Ascending	Descending	Increasingly Variable
1 vs. 2	2.93*	1.56	1.75
1 vs. 3	1.58	1.52	4.64**
1 vs. 4	1.60	2.64*	1.79
2 vs. 3	1.85	2.37*	2.66*
2 vs. 4	4.69**	4.12**	1.02
3 vs. 4	2.54*	1.74	2.60*

\*\* Significant beyond .01 point.

\* Significant beyond .05 point.

All three of the programs produced significant changes in variability during the hour. Within the V+ Program, for example, the subjects were significantly more consistent (i.e., less variable) during the third quarter than in the other three. But within the D Program, the significant effects appear at different times: subjects were more consistent during the first two quarters, becoming significantly more variable during the latter two quarters. The A Program follows a still-different pattern, and the subjects were significantly more consistent during the first and last quarters than during the middle ones. With each program, half of the six possible F tests proved statistically significant.

These results show that the three programs had quite different effects on the subjects, even though all three programs used the same 23 musical selections. Not only did all three programs produce significant changes in variability during the hour, but the patterns of variability were characteristically different for each program.

The music program's powerful effects on variability may shed light on some of the apparent conflicts in the literature. For one thing, significant changes in variability might overshadow differences between means, leading investigators to conclude that groups do not differ significantly when they really do. At the same time, it seems conceivable that an investigator who is not looking for changes in variability might well overlook them.

Table 8 shows the  $F$  ratios comparing the three programs for each quarter hour. Once again, exactly half of the comparisons reached statistical significance, all but one of them at the .01 probability point. These significant  $F$  values show that each program had significantly different variability than the others in at least one quarter hour. Moreover, they underline the similarity between the A and V+ Programs, which have similar variabilities except during the third quarter hour. On the other hand, A and D differ in variability during three of the four quarter hours, while V+ and D differ significantly during two of the four. Both A and V+ differ significantly from D during the last two quarters; this finding may indicate that the stimulation progression has a cumulative effect.

TABLE 8

$F$  Ratios Comparing Variabilities Between Programs, by Quarters

Quarter Hour	Comparison		
	A vs. D	A vs. V+	D vs. V+
1	3.41**	2.21	1.55
2	1.33	2.32	1.73
3	3.28*	3.33**	10.93**
4	14.49**	1.98	7.31**

\*\* Significant beyond .01 point.

\* Significant beyond .05 point.

It should be noted that, because the programs affected variability so strongly, there is no really meaningful way to test the significance of the differences between mean response times. Most of the orthodox methods for comparing central tendencies presuppose equal variabilities. While assorted transformations of scale can sometimes equalize variabilities, none of the common transformations seems suited to coping with the present data.

## Longest Response Times

A somewhat different approach is segregating each subject's longest response time. Table 9 gives mean "longest" responses for the programs and shows their distribution by quarter hours. The frequencies have not been analyzed statistically, since some cell entries are too small for computing chi-square. Nevertheless, there appear to be differences among the programs. The V+ Program concentrates the majority of longest responses in the first quarter. On the other hand, the A Program gives more longest responses during the last two quarters. Longest responses are more or less evenly scattered with the D Program, although they seem to increase as the hour wears on.

TABLE 9  
Longest Responses:  
Means, Variances, and Distribution by Quarters

Quarter	Programs		
	Ascending	Descending	Increasingly Variable
1	4	5	12
2	3	4	3
3	7	5	1
4	7	7	5
Total	21	21	21
Mean	3.926	5.519	4.847
Variance	6.082	12.685	5.741

The A Program has the shortest "mean longest" response, and the D Program the longest. Just as with the quarter-hour means, these "longest" responses show gross differences in variability which make it difficult to compare them statistically. Obviously, the Descending Program gave much more variable "longest" responses: significantly more variable than the V+ Program ( $F = 2.21$ ,  $p < .05$ ), and almost significantly more ( $F = 2.09$ ,  $F_{.05} = 2.12$ ) than the A Program. Clearly, the Descending Program made "longest" responses more variable.

The largest mean difference -- Ascending vs. Descending -- reaches statistical significance ( $t = 1.684$ ,  $p < .05$ ), although the other differences do not. Thus the very fact that any of these comparisons shows significance, despite the wide differences in variability, attests to the powerful effect the programs had on the subjects' performance.

### More-Stimulating vs. Less-Stimulating Selections

A more specific comparison may be made within the Increasingly Variable Program, since it presented more-stimulating and less-stimulating selections alternately. If the ranked stimulation values of the selections actually affect the subjects' response times, subjects should give faster responses with more-stimulating music, and slower ones with less-stimulating music. Furthermore, since the contrast between adjacent selections grows progressively greater, so should this difference become more pronounced later in the hour.

Testing this prediction requires four mean response times for each of the 21 subjects in the V+ group:

1. Mean for less-stimulating selections during the first half hour (trials 1, 2, 3, and 6): 1.324 seconds.
2. Mean for less-stimulating selections during the second half hour (trials 9, 10, 12, 15, and 17): 1.341 seconds.
3. Mean for more-stimulating selections during the first half hour (trials 4, 5, 7, and 8): 1.424 seconds.
4. Mean for more-stimulating selections during the second half hour (trials 11, 13, 14, and 16): 1.064 seconds.

While these means do not balance the serial positions of trials perfectly, they do allow at least rough comparisons. During the first half hour, the more-stimulating selections gave a slightly longer response time than the less-stimulating selections. However, mean response time was considerably shorter for the more-stimulating selections during the second half hour, while it remained about the same for the less-stimulating selections.

Since the variabilities about these four group means were similar, the differences between means may be evaluated with Fisher's  $t$  for paired measures. Response-time means did not differ significantly during the first half hour. During the second half hour, however, the subjects responded significantly faster when listening to the more-stimulating selections ( $t = 2.451$ ,  $p < .025$ ).

To clarify just what happened, we may ask whether the subjects took progressively longer to respond during less-stimulating selections, or responded more promptly during more-stimulating selections, or both. The answer is clear-cut. Mean response time for less-stimulating selections was very slightly longer during the second half hour, but this difference is not statistically reliable. With more-stimulating music, on the other hand, response times were very significantly faster -- more than a third of a second less -- during the second half hour ( $t = 2.857$ ,  $p < .005$ ). Thus less-stimulating music did not reduce the subjects' alertness; more-stimulating music improved their performance very significantly.

#### Consistency (Reliability) of Individual Performance

The remaining criterion is individual consistency during the hour, which can be obtained conveniently by averaging each subject's response times during the first two and the last two quarters, then correlating them. Table 10 gives rhos for the three programs.

TABLE 10

#### Correlation Between Response Times, by Half Hours

Program	rho
Ascending	+.713**
Descending	+.689**
Increasingly Variable	-.082

\*\* Significant beyond .01 level.

Both Ascending and Descending Programs gave very significant positive correlations between means for the two half hours; interestingly, the two coefficients are roughly equal. But the V+ Program yielded a slight negative correlation, which is not significantly different from zero. While these findings can be summarized in a few words, they have crucial and far-reaching import for vigilance with music.

These correlations tell us that either gradual increase or gradual decrease in the music's stimulation value affects the individual consistently and predictably. The man who is fastest during one part of the hour will still respond quickly during other parts of the hour. The man who is slow during one part of the hour will tend to be slow during other parts of the hour.

However, the Increasingly Variable Program completely shatters this individual consistency. If a man responds promptly during part of the hour, we can make no useful prediction about how he will respond at other times -- he may respond quickly, he may respond slowly, or he may be merely average. This finding demonstrates that the contrast between adjacent selections is a potent -- and potentially dangerous -- force in programming music. It suggests that marked contrasts should be used very cautiously, because they affect response times unpredictably.

As a source of hypotheses for further experiments, the consistency of individual performance was also examined by quarter hours (Table 11). The rhos for the V+ Program show conclusively that contrast does affect subjects differentially: subjects do not perform the same way with low contrast as they do with higher contrast, and only two of the quarters are correlated significantly. Subjects respond more consistently with the Descending Program, although the first quarter hour has no significant correlation with the other three quarters. The Ascending Program is still more consistent, and four of the six rhos are significant. With A, quarters near each other tend to be related more closely. No quarter is unrelated to all the rest, as the first quarter is with D and V+.

TABLE 11

## Correlation Between Response Times, by Quarter Hours

Quarter Hours Compared	Programs		
	Ascending	Descending	Increasingly Variable
1 vs. 2	+.498*	+.096	-.109
1 vs. 3	+.510*	+.064	-.103
1 vs. 4	+.196	+.144	-.181
2 vs. 3	+.547**	+.506*	+.246
2 vs. 4	+.515*	+.534**	+.244
3 vs. 4	+.113	+.513*	+.465*

\*\* Significant beyond .01 level.

\* Significant beyond .05 level.

## Attitude -Scale Scores

The Thurstone-type attitude scale, already shown in Table 4, was included primarily to check whether the three experimental groups had comparable attitudes toward the music used in the programs. We cannot merely assume that attitudes were equivalent; because the attitude variable is potentially so important, it must be measured when it cannot be controlled. Suppose, for instance, that one group liked the music very much, while another group disliked it very much. These two groups would probably perform quite differently -- not because of variations in the program sequence, but because they had significantly different attitudes toward music even before the experiment began.

The mean attitude scores for the Ascending, Descending, and Increasingly Variable groups were, respectively, 5.08, 5.48, and 5.22. On the seven-point scale, these means clustered between "slightly favorable" (5.00) and "favorable" (6.00). These means suggest that, if anything, the Descending group -- with its very slightly more favorable attitude -- might have responded a bit more favorably to music than the other groups did. However, the differences among these means are not statistically significant, which indicates that the three groups had essentially comparable attitudes toward music. Thus the observed differences among the groups cannot be attributed to differing attitudes toward music.

## Conclusions

Altogether, these results indicate that the programs caused gross differences in response measures. These findings strongly support O'Neill's theoretical position. Furthermore, since the three program sequences were based on O'Neill's ranked stimulation values, these positive findings indirectly confirm the validity of the rankings themselves.

In terms of the three criteria which were established, the Ascending Program gave better performance than the Descending and Increasingly Variable Programs. The A Program's mean response times for quarter hours were always faster than those for the D Program, and about as fast as with the V+ Program. "Mean longest" response time was also fastest with the A Program, and its very significant consistency (or reliability) coefficient demonstrates that it affected subjects consistently and predictably. Further studies of vigilance with functional music should examine variations of this simple ascending sequence to see if it can be improved further, and to establish whether slight to moderate contrast between adjacent selections would give still better performance.

## SUMMARY

1. This experiment attempted to answer two questions: (a) Do the ranked stimulation values of musical selections affect response times on a vigilance task, and (b) Does the sequence (order) of selections affect performance?
2. Sixty-three enlisted subjects worked at the task for one hour while listening to three sequences of the same 23 musical selections. The Ascending (A) sequence began with the least stimulating selection and grew steadily more stimulating. The Descending (D) sequence used exactly the opposite order. The Increasingly Variable (V+) sequence began in the middle of the rank order and, maintaining roughly the same average stimulation, increased the contrast between adjacent selections.
3. When response times were grouped by quarter hours and averaged, the A Program gave consistently faster responses than the D Program. The A Program gave faster times than the V+ Program during the first and last quarters.
4. Order of selections had extensive and statistically significant effects on variability. Variances differed widely, not only from program to program, but within each of the programs. All three programs produced significant changes in variability during the hour, and the patterns of variability were characteristically different for each. Such effects, if overlooked, could account for some of the negative results and conflicting results in the literature.
5. Considering only each subject's longest response time during the hour, means were shortest for A and longest for D; this difference was statistically significant. In addition, the distribution of longest responses by quarters appeared to depend on the programs.
6. During the second half hour of the V+ Program, subjects responded significantly faster with more-stimulating selections than with less-stimulating ones -- and very significantly faster than during the first half hour.
7. Both of the gradually changing programs (A and D) gave very significant reliability coefficients, indicating that subjects responded consistently and predictably during the hour. With the V+ Program, however, the reliability coefficient fell to essentially zero. Thus marked contrasts between adjacent selections destroyed the consistency of individual performance.
8. In terms of all of these criteria -- mean response times, variability, and individual consistency -- the Ascending Program was superior to both Descending and Increasingly Variable Programs.
9. These results validate O'Neill's theoretical position by demonstrating that the effects music has on performance depend on both (a) the ranked stimulation values of the musical selections and (b) the program, or sequence of selections.

## REFERENCES

1. Cardinell, R. L. Music in industry. In Schullian, D. M., & Schoen, M., Music and medicine, pp. 352-366. New York: Henry Schuman, 1948.
2. Cardinell, R. L. & Burris-Meyer, H. Music in industry today. J. acoust. Soc. Amer., 1947, 19, 547-549.
3. Cattell, R. B. & Anderson, Joan C. The measurement of personality and behavior disorders by the I.P.A.T. Music Preference Test. J. appl. Psychol., 1953, 37, 446-454.
4. Cattell, R. B. & Saunders, D. R. Musical preferences and personality diagnosis: I. A factorization of one hundred and twenty themes. J. soc. Psychol., 1954, 39, 3-24.
5. Ellis, D. S. & Brighthouse, G. Effects of music on respiration and heart rate. Amer. J. Psychol., 1952, 65, 39-47.
6. Freeburne, C. M. & Fleischer, M. S. The effect of music distraction upon reading rate and comprehension. J. educ. Psychol., 1952, 43, 101-109.
7. Henkin, R. I. A factorial study of the components of music. J. Psychol., 1955, 39, 161-181.
8. Henkin, R. I. A reevaluation of a factorial study of the components of music. J. Psychol., 1957, 43, 301-306.
9. Henkin, R. I. The prediction of behavior response patterns to music. J. Psychol., 1957, 44, 111-127.
10. Nelson, D. O. & Finch, L. W. Effect of audio-analgesia on gross motor performance involving acute fatigue. Res. Quart. Amer. Assoc. Hlth. Phys. Educ. Recr., 1963, 33, 588-592.
11. Offer, D. & Stine, Diane. Function of music in spontaneous art productions. Arch. gen. Psychiat., 1960, 3, 490-503.
12. O'Hanlon, J., Jr., Schmidt, E. A., & Baker, C. H. A study to determine the effects of placebos upon performance in a vigilance task. Technical Report 750-3, Human Factors Research, Inc., Los Angeles, Calif., January 1964.
13. O'Neill, D. M. Music to enhance the work environment. Mgmt. of Pers. Quart., 1966, 5, 17-23.

14. Shatin, L., Lussier, L. W., & Kotter, W. Fear in ECT and its alleviation through music. Psychiat. Quart. Suppl., 1958, 32, 200-223.
15. Skelly, C. G. & Haslerud, G. M. Music and the general activity of apathetic schizophrenics. J. abnorm. soc. Psychol., 1952, 47, 188-192.
16. Traxel, W. & Wrede, G. Hautwiderstandsänderungen bei Musikdarbietung. (Changes in GSR while listening to music.) Z. exp. angew. Psychol., 1959, 6, 293-309.
17. Van Krevelen, Alice. The influence of music on reactions to frustration. J. gen. Psychol., 1963, 68, 107-110.
18. Weidenfeller, E. W. & Zimny, G. H. Effects of music upon GSR of depressives and schizophrenics. J. abnorm. soc. Psychol., 1962, 64, 307-312.
19. Wokoun, W. Vigilance with background music. Technical Memorandum 16-63, U. S. Army Human Engineering Laboratories, Aberdeen Proving Ground, Md., August 1963.
20. Wyatt, S. & Langdon, J. N. Fatigue and boredom in repetitive work. Report No. 77, Industrial Health Research Board, London, England, 1937.
21. Zimny, G. H. & Weidenfeller, E. W. Effects of music upon GSR and heart-rate. Amer. J. Psychol., 1963, 76, 311-314.

Unclassified

Security Classification

**DOCUMENT CONTROL DATA - R & D**

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

<b>1. ORIGINATING ACTIVITY (Corporate author)</b> U. S Army Human Engineering Laboratories Aberdeen Proving Ground, Maryland 21005		<b>2a. REPORT SECURITY CLASSIFICATION</b> Unclassified	
		<b>2b. GROUP</b>	
<b>3. REPORT TITLE</b>  EFFECTS OF MUSIC ON WORK PERFORMANCE			
<b>4. DESCRIPTIVE NOTES (Type of report and inclusive dates)</b>			
<b>5. AUTHOR(S) (First name, middle initial, last name)</b>  William Wokoun			
<b>6. REPORT DATE</b> January 1968		<b>7a. TOTAL NO. OF PAGES</b> 44	<b>7b. NO. OF REFS</b> 21
<b>8a. CONTRACT OR GRANT NO.</b>		<b>8b. ORIGINATOR'S REPORT NUMBER(S)</b>  Technical Memorandum 1-68	
<b>b. PROJECT NO.</b>		<b>8c. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)</b>	
<b>c.</b>			
<b>d.</b>			
<b>10. DISTRIBUTION STATEMENT</b>  This document has been approved for public release and sale; its distribution is unlimited. Released to Department of Commerce for sale to the public.			
<b>11. SUPPLEMENTARY NOTES</b>		<b>12. SPONSORING MILITARY ACTIVITY</b>	
<b>13. ABSTRACT</b>  Sixty-three subjects worked at a vigilance task for an hour while listening to one of three musical programs. While all three programs included the same 23 selections, the Ascending Program grew steadily more lively, and the Descending Program grew steadily less lively, while the Increasingly Variable Program progressively increased the contrast between adjacent selections. Results showed that changing the sequence of the 23 selections profoundly affected reaction times and variabilities, as well as individual consistency. Subjects performed the vigilance task better with the Ascending Program than with the other two programs.			

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Human Factors Engineering Music Work Performance Vigilance						