Auditory Signal Detectability in Psychiatric Patients and Non-patients

Gerard E. Bruder, Samuel Sutton, Barry J. Gurland and Allan Yozawitz

Biometrics Research, New York State Department of Mental Hygiene

Acknowledgments

This research was performed under Grant No. MH 18422 from the National Institute of Mental Health, United States Public Health Service. The authors wish to acknowledge the aid of Robert Laupheiner and Raymond Simon for the design and maintenance of the equipment, and to thank Dr. M. Wallach, Director of Brooklyn State Hospital, for the use of Brooklyn State Hospital facilities. The authors also benefitted from the advice of Drs. Harvey Babko and Joseph Fleiss.

(Enclosure #1)
Until recently, there was little evidence to contest statements that there are no differences in the auditory thresholds of schizophrenic patients and normals (e.g., Yahei, 1966). Studies using classical psycho-physical techniques to measure thresholds of schizophrenics and normals generally agreed in showing no differences (Bartlett, 1935; Ludwig, Wood, & Downs, 1962; Rappaport & Hopkins, 1969). The only early findings of higher auditory thresholds in schizophrenics were obtained using conditions designed to produce "day-dreaming" (L. E. Travis, 1924; R. C. Travis, 1926).

More recent studies, which have used not only classical threshold measures (Levine & Whitney, 1970) but also signal detectability measures (Emmerich & Levine, 1970; Rappaport, Hopkins, Silverman, & Hall, 1972) have, in contrast, found elevated thresholds and lower signal detectability in schizophrenic patients. The chronic schizophrenics tested by Levine and Whitney had an 8 dB higher average threshold than normal controls. The authors do, however, point out that the higher thresholds of the schizophrenics could have resulted from either lower sensitivity (d') or stricter response criterion. In a further study, Emmerich and Levine used a forced-choice signal detection procedure that provides a criterion-free measure of sensitivity, and they also found 8 dB higher thresholds (i.e., dB needed for 85 percent correct performance) in chronic schizophrenics as compared to normal controls. Rappaport et al. used a "Yes vs. No" signal detection task and similarly found paranoid schizophrenics to have generally lower sensitivity (d') than normals.

Levine & Whitney (1970) have interpreted the conflicting reports of the presence vs. absence of schizophrenic-normal differences in auditory
Method

Subjects

The 36 patients were selected from the new admissions at Brooklyn State Hospital and were tested within two weeks after admission. Selection was restricted to include only cooperative, male patients between the ages of 18-30 years, who were not known to be mentally retarded, brain-damaged or addicted to drugs or alcohol. The mean age of the patients was 23.4 years. The 50 non-patients were male college students and referrals from New York State employment service, who were between the ages of 18-31 years. The average age of the non-patients was 20.6 years.

Apparatus

The clicks were presented by the right phone of Sharpe Mark II circum-aural earphones. The clicks were generated by a negative-going electrical pulse with an exponential return to base. The time constant of the exponential return to base (1/e x peak amplitude) was 0.1 msec. An attenuator (Hewlett-Packard 5-W 600-Ω) permitted attenuation of the click intensity in 1 dB steps. Further details of the apparatus are described elsewhere (Babkoff & Sutton, 1968).

Procedure

Forced-choice threshold. We used a three-interval temporal forced-choice technique that consists of presenting a click signal in one of three observation intervals and having the subject (s) indicate which of the intervals contained the signal. This technique has a number of advantages that dictated its use in this study. First, it eliminates the response criterion problem, which can be especially troublesome in patient-normal comparisons
(Clark, Brown, & Rutschmann, 1967). Second, the experimenter knows when the S is correct or incorrect. By presenting a signal that is clearly audible, it is possible to test whether or not the S understands the task and is performing correctly. This is clearly important when testing psychiatric patients. Third, since the S is forced to respond on each trial, he must maintain at least a minimum attention to his task. Fourth, in a three-interval forced-choice task the subject may respond based only on his perception of a difference in one of the three intervals. We need not be dependent on his memory of the signal from prior trials, as is the case in a two-interval forced-choice and "Yes vs. No" detection task. This may reduce the influence of a non-sensory memory factor on performance in the threshold task.

In order to provide efficient estimates of forced-choice thresholds, we used an "up-down" staircase procedure (see Levitt, 1971, for the rationale underlying such procedures). Campbell (1963) introduced a Block Up and Down, Two-Interval, Forced-choice (BUDTIF) procedure and he obtained data suggesting that, "...in spite of the essential naiveté and lack of selection of the subjects used, the BUDTIF method quickly provided information comparable in precision and magnitude to that obtained by utilizing the most exhaustive experimental procedure in general use" (p. 1733). The BUDTIF procedure might, therefore, be an especially valuable tool for measuring thresholds of patients in as brief a period possible. We adopted the BUDTIF procedure to permit the estimation of the click intensity needed for 67% correct responses (50% adjusted for chance) in the three-interval forced-choice task. The specific details of this procedure are outlined below.
(1) On each trial, a signal was presented in one of three observation intervals and the S indicated which interval contained the signal.

(2) A constant signal intensity was used during each block of three trials.

(3) The signal level used in a block depends on the S's performance in the prior block. The "stepping rules" for estimating the 67% correct point are as follows:

0 or 1 correct — increase the level by one step
2 correct — stay at the same level
3 correct — decrease the level by one step

(4) We used both an "up" and a "down" staircase, which were randomized across blocks. The staircases were started at signal levels known to be well below ("up" staircase) or above ("down" staircase) the target level. The advantages of using such a random-double staircase are discussed by Cornevelet (1962).

(5) We started the staircases using a step size of 3 dB and reduced the step size to 1 dB when the "up" and "down" staircases converged.

(6) The "up" and "down" staircases were run for a combined total of about 50 blocks, with some exceptions where the staircases were terminated earlier (MIN = 30 blocks) or extended beyond this point (MAX = 98 blocks). The average number of blocks for patients was 48 and for non-patients was 54.1

(7) An estimate of the signal level yielding 67% correct responses is given by the median of the signal levels revisited in the "up" and "down" staircases.
Figure 1 shows staircase data for one patient and one non-patient. Note first that the initial portion of the "down" staircase drops toward lower signal levels showing that the Ss were 100% correct at the higher signal levels, and they therefore were performing properly in the task. No staircase was begun unless the S was 100% correct at the highest signal level. Note secondly that the "up" and "down" staircases converge and remain around the same signal levels. This is also a sign that the Ss were performing properly in the threshold task (see Corssweet, 1962).

Thresholds were measured in a sound-attenuating booth. There was a low, constant ambient noise produced by a fan in the booth. The circumaural earphones attenuated this ambient noise to a point that was barely audible. Each S was seated in the booth and given instructions for the three-interval forced-choice task. A few practice trials were then run at a highly audible signal level to check whether or not the S understood the instructions. If the S was not 100% correct in these practice trials, he was given further instructions and practice. Several patients were unable to perform correctly in the task and were not used in the study. After instructions and practice were successfully completed, the staircases were begun. The amount of time needed to complete testing ranged from about 45-90 min. depending on the number of blocks. A brief rest outside the booth was given after about the first 20 min. of testing, and additional rests were given during longer staircases or upon the request of the S.

Each trial consisted of the following sequence of events. At the beginning of each trial a rod flash signaled the S that we were ready to begin. When the S was ready he initiated the trial by pressing a telegraph
Fig. 1. Block Up and Down, Three-Interval, Forced-choice staircase data for a patient and a non-patient.
key, which produced another red flash signaling him that the trial had begun. After 3.6 sec., there were three red flashes each of 0.5 sec. duration and separated by 0.7 sec. These three flashes served as the observation intervals. The click signal was presented 0.1 sec. after the onset of one of the observation intervals. The S indicated his choice by pressing one of three response buttons. If he was correct in his choice, a blue "feedback" light went on when he pressed the correct button. No feedback was given if he was incorrect. Ss were instructed that they would earn a one cent reward each time they were correct and the blue light went on. Patients were also given a bonus for performing in the task. The non-patients received the one cent rewards in addition to their hourly wages.²

Interview. After threshold measurement (usually on the same or following day) each S was interviewed using a structured mental state interview schedule that was adapted from the Present State Examination (Wing, Birley, Cooper, Graham & Isaacs, 1967) and from the Psychiatric Status Schedule (Spitzer, Fleiss, Burdock, & Haráesty, 1964). The origin of this schedule is described by Cooper et al. (1969). Ratings from the interview provided the information for obtaining scores on 20 factors of psychopathology that were derived in a prior factor analysis of the ratings of 500 patients from Brooklyn and London (Fleiss, Gurland, & Cooper, 1971). A profile showing the standard scores ($\bar{X} = 50, Sd = 10$) on each factor of psychopathology was thereby obtained for each patient and non-patient.³ Patients and non-patients were divided into sub-groups based on their profile scores using a classification scheme previously utilized in a study by Gurland et al. (1970). The classification scheme can be illustrated with reference to Figure 2, which shows the average profiles for sub-groups that contained a minimum of 5 Ss.
Fig. 2. Average symptom profiles for patient and non-patient sub-groups.
Patients and non-patients were assigned into sub-groups based on their profile scores in three major areas of psychopathology: (1) Mood disturbance, which consists of factors usually associated with a diagnosis of depression; (2) Hypomania, which is associated with a diagnosis of schizophrenia and manic-depressive manic; (3) Conceptual or perceptual disorganization, which consists of factors most often associated with a diagnosis of schizophrenia. A standard score higher than 60 (i.e., higher than 1 Sd above the mean) on any one factor in one of these areas of psychopathology was sufficient for classification in that area. Thus, a S in the Disorganized sub-group scored high on at least one of the disorganization factors, but not on factors in the other areas. A S in the Mood-Disorganized sub-group scored high on at least one factor in both the mood and disorganization areas. The remaining groups were formed in the same manner.

Results

The frequency distributions of thresholds for all patients and non-patients are shown in Figure 3. The mean threshold for patients was 3 dB higher than the mean threshold for non-patients. This difference was significant at the .01 level ($t = 3.06, d.f. = 84$). Note that only 6 percent of the non-patients had thresholds higher than -70 dB attenuation, while 14% of the patients had thresholds above this value. The distribution of thresholds for patients also shows more dispersion ($Sd = 5.98$) than the distribution for non-patients ($Sd = 3.49$).

It is, however, clear from the data in Figure 3 that not all patients showed elevated thresholds. About half of the patients fell in the "normal"
Fig. 3. Frequency distributions of forced-choice thresholds for patients and non-patients.
non-patient threshold range. In an effort to identify the sub-group or sub-groups that show elevated thresholds, we compared the thresholds of patients and non-patients classified into the five sub-groups whose profiles are shown in Figure 2. Thirty-nine non-patients had low scores on all the factors of psychopathology shown in Figure 2. Interestingly, eight non-patients showed high scores on disorganization factors, and had the highest average score of all sub-groups on the visual hallucination factor. Ten Disorganized patients showed a similar average profile, but had a higher score on incomprehensibility. Thirteen Non-Disorganized patients had high scores on a number of factors in both these areas, and had higher scores than the other sub-groups on the factors of depression, retarded speech, auditory hallucinations, and flat affect. Five Manic-Disorganized patients also showed very high scores on a number of factors, including hypomania, lack of insight, the delusions, and incomprehensibility. The remaining patients and non-patients were classified into sub-groups with three or less Ss, and their data are not considered in further analyses.

The means, standard deviations, and distributions of thresholds for the five sub-groups are given in Figure 4. A one-way analysis of variance between the thresholds for the 5 sub-groups was found to be significant at the .05 level ($F = 3.44$, d.f. = 4,70). However, Scheffe's multiple comparison tests ($\alpha = .05$) revealed that only the threshold difference between Non-Disorganized patients and non-patients with low profile scores was significant. The thresholds of Disorganized patients and Disorganized non-patients did not differ from the thresholds of non-patients with low profile scores.
Fig. 4. Frequency distributions of forced-choice thresholds for sub-groups of patients and non-patients.
In contrast, the distribution of thresholds for Moody-Disorganized patients was displaced toward higher threshold values. The mean threshold for Moody-Disorganized patients was 5 dB higher than the mean threshold for non-patients. The Manic-Disorganized patients also showed a higher mean threshold than the non-patients, but the difference was less than found for the Moody-Disorganized patients, and it was not significant.

The form of the "up-down" staircases for patients and non-patients in the different sub-groups were very similar. Even patients with the highest thresholds (e.g., the patient in Figure 1) showed regular "up-down" staircases that differed from the non-patient staircases only with respect to the signal levels needed for 67% correct performance. For example, the "up" and "down" staircases for patients and non-patients in the different sub-groups converged in about the same number of blocks. Moreover, the staircases of patients and non-patients in the different sub-groups showed about the same variability. An index of the variability of the staircases is given by the semi-interquartile range of signal levels revisited in the "up" and "down" staircases. A one-way analysis of variance between the semi-interquartile ranges for the 5 sub-groups proved not to be significant at the .05 level ($F = 1.45$, d.f. = 4,70). The average semi-interquartile ranges of the staircases for the sub-groups were from 0.7 - 1.4 dB. Thus, the patient sub-groups were not more variable in their staircase detection performance than the non-patient sub-groups.

Discussion

Patients and non-patients scoring high on only conceptual or perceptual disorganization factors, which are usually associated with a hospital
diagnosis of schizophrenia (Gurland et al., 1970), did not have higher thresholds than non-patients with low profile scores. However, patients with high scores on both disorganization factors and mood factors (i.e., depression and retarded speech) did show elevated thresholds. These Mood-Disorganized patients required 5 dB more acoustic energy to obtain the same level of signal detectability (d') as non-patients with low profile scores. Gurland et al. (1970) have found that patients with high scores on both mood and disorganization factors are diagnosed in New York hospitals as schizophrenic in about 80% of the cases, while in London hospitals they are diagnosed as depressive in about 80% of the cases. This is an example of a difference in diagnostic criteria that could dramatically influence whether or not elevated auditory thresholds are obtained in a sample diagnosed as schizophrenic. It is our hope that measures such as auditory signal detectability will contribute toward the development of objective culture-free indicators of psychopathology (Zubin, 1969).

The very high profile scores of Mood-Disorganized patients on the symptom factors of depression, retarded speech, auditory hallucinations, and flat affect appear to best differentiate them from patients and non-patients in the other sub-groups. Although the elevated thresholds of the Mood-Disorganized patients may be related to their high scores on one or more of these symptoms, it may be argued that their elevated thresholds are more reflective of the overall severity of their symptomatology. An index of overall severity was, therefore, obtained by computing the average score each S received on all 22 factors of psychopathology in the symptom profiles. There was no correlation between the patients' average profile
scores and their thresholds \( r = .02 \). Nor did the Moody-Disorganized patients show significantly higher average profile scores than the Disorganized and Manic-Disorganized patients (Scheffe test, \( \alpha = .05 \)). The patients' average profile scores in each area of psychopathology (i.e., Moody, Disorganization, Hypomania) were also determined, and the only area of psychopathology where Moody-Disorganized patients scored significantly higher than Disorganized and Manic-Disorganized patients was in the area of Moody (\( \alpha = .05 \)). The overall severity of symptomatology does not, therefore, appear to be a crucial determinant of the Moody-Disorganized patients' elevated thresholds.

Since all patients were receiving medication from about the time of hospital admission, it could be argued that the threshold difference between Moody-Disorganized patients and non-patients was due to drug effects. The types and dosages of the drugs received by patients in the different subgroups were similar, with the exception of two types of phenothiazines — Thorazine (chlorpromazine) and Stelazine (trifluoperazine HCL). In the Moody-Disorganized sub-group, there were a higher proportion of patients taking Stelazine but a lower proportion taking Thorazine as compared to the Disorganized sub-group. The Manic-Disorganized sub-group showed a relatively high proportion of use of Stelazine and Thorazine. Statistical analysis, however, revealed no significant effects of the presence vs. absence of Thorazine or Stelazine on thresholds (\( p > .05 \)). Nor were there significant correlations between dosage level of Thorazine or Stelazine and thresholds (\( p > .05 \)). Moreover, previous studies have not found phenothiazine drugs to have any significant effects on auditory thresholds (Ludwig et al.,
1962; Rappaport & Hopkins, 1969). Rappaport et al. (1972) did report that phenothiazines differentially affect auditory signal detectability in paranoid and nonparanoid schizophrenics, but the detectability differences they report are very small (equivalent to 1 dB or less difference) and are not statistically significant (Emmerich & Levine, 1973). It would, therefore, be difficult to account for the elevated thresholds of a single sub-group, i.e., Moody-Disorganized patients, based solely on drug effects.

One major question that remains to be answered is whether or not the patient vs. non-patient differences in auditory signal detectability demonstrated in this study and in prior studies (Emmerich & Levine, 1970; Rappaport et al., 1972) reflect real sensory differences between the patients and non-patients. Although the use of signal detectability techniques eliminates possible confounding due to non-sensory response criterion differences, this technique does not insure that detectability differences are sensory differences. Factors such as attention and motivation are still free to be confounded with detectability differences. We have attempted to design our study to help overcome possible "poorer" attention and motivation in patients. For example, having the patient initiate each trial when he was ready to begin, and using a forced-choice task, where the patient must respond on each trial, aid in maintaining attention. Also, the "trial to trial" feedback and rewards for correct performance aid in maintaining motivation. An index of how well the patients did perform is given in their "up" and "down" staircase data. The patients' staircases were just as good as those for non-patients, as was demonstrated, for example, by the equal variability of the patient and non-patient staircases. An advantage of the staircase procedure used in this study is that it keeps the performance of patients and non-patients at the
same level of correct detections (i.e., at the target level of, in this case, 67% correct).

If the patient vs. non-patient detection difference does result from an auditory deficit in the patients, it should be possible to trace the origins of this deficit in the auditory system. Audiologists have a variety of tests that could aid in establishing the locus of a hearing loss in patients. One such audiological test that could prove to be of value is a test for recruitment. Although no audiological measures were obtained for patients showing elevated thresholds in our study, i.e., the Hoody-Dysorganized patients, Levine and Whitney (1970) found that chronic schizophrenics showed not only higher detection thresholds, but also lower thresholds of unpleasantness to intense tones. They suggest that this may reflect an abnormally rapid rise of loudness in chronic schizophrenics similar to that encountered in recruitment. They also indicate that the presence of recruitment would tend to localize the site of the deficit to the cochlea.

A second test that can be used to find the locus of a hearing deficit is "brief tone audiometry", which is also referred to as auditory temporal integration. This measures the ability of the auditory system to utilize or integrate the acoustic energy in tones. Specifically, what is measured is the thresholds to tones of various durations, ranging from very brief (e.g., 0.01 sec.) to much longer tones (e.g., 1 sec.). The resulting auditory integration function has been shown to be altered in cases of damage to the cochlea (e.g., Cengel & Watson, 1971; Harris, Haines, & Myers, 1958; Miskoczy-Fodor, 1953; Pederson & Elberling, 1973; Sanders & Honig, 1967). Moreover, the auditory integration function is altered in an opposite manner
in cases of brain damage in the auditory cortical regions (Gersuni, Baru, Karaseva, & Tonkonogii, 1971). If, as some have suggested (Emmerich & Levine, 1970; Meñick, 1969), the development of "schizophrenia" is related to brain damage resulting from birth complications such as asphyxia, then patients exhibiting elevated auditory thresholds may show auditory integration functions that more closely resemble those found in cases of brain damage to auditory cortex. One specific prediction, suggested by the auditory integration functions for brain damaged patients (see Gersuni et al., 1971), is that patients with hearing losses resulting from an auditory deficit that has cortical origins will show elevated thresholds for very brief tones, but not for long tones. We plan to test predictions such as these to help establish the basis of the elevated thresholds that have been found for subgroups of psychiatric patients in this study and in previous work.
References


Mednick, S. A. Breakdown in individuals at high risk for schizophrenia: behavioral and automatic characteristics and possible role of perinatal complications. An address delivered to the Society for Research in Child Development, Santa Monica, 1969.


Footnotes

1. Although the non-patients' staircases were on the average run a few more blocks than the patients' staircases, the non-patients' thresholds did not benefit from these extra trials. Thus, there was no correlation between the number of blocks non-patients were run and their thresholds. Also, the mean threshold of non-patients run more than 50 blocks did not differ from the mean threshold for non-patients run up to 50 blocks.

2. The first 9 patients and 14 non-patients tested were given "feedback" flashes when correct, but they did not receive the one cent rewards, nor did the patients receive a bonus. The difference in the average thresholds for these patients and non-patients was found to be exactly the same as for those who received rewards. The data for the unrewarded and rewarded Ss were therefore pooled.

3. The standardization of factor scores was based on the ratings of the 500 patients in Brooklyn and London (see Fleiss et al., 1971). No standardization on non-patients is available at this time.

4. Nineteen of these non-patients did score high on one or more factors not utilized in the assignment of Ss into sub-groups (i.e., anxiety, somatic concern, observed belligerence, reported belligerence, obsession, disorientation, and depersonalization). These non-patients were found to have exactly the same mean threshold as non-patients who scored low on all factors.