Response Delay, Psychophysiologic Activation, and Recognition of One's Own Voice

GUILLERMO OLIVOS, M.D.*

One hundred and ten subjects were confronted with the verbal discriminatory task of differentiating their own voices from those of other people. The Ss were asked to delay their verbal report for 12 sec., while GSR, GSP, and EEG were recorded. A significantly larger number of Ss in the experiment recognized their own voices as compared to the results obtained by other experimenters who requested an immediate verbal report from their Ss. A year later 65 of the original 110 Ss were confronted with the same task but requested to give their report immediately after listening to the recorded voices. The results supported the hypothesis that allowing time for the autonomic reaction to reach a maximum improves the Ss' identification of their own voices. A significantly different physiological response was observed when the Ss listened to their own voices as compared to listening to the voices of other people.

In a previous study Rousey and Holzman[4] showed that 62% of Ss failed to recognize their own tape-recorded voices. In that study Ss were only moderately familiar with listening to their own recorded voices. In another study Holzman *et al.*[2] demonstrated that Ss showed significantly higher psychophysiological activation by their own voices than by the voices of others, whether or not Ss were aware that they were hearing their own voices.

Reiser and Block[3] have demonstrated that delaying the S's verbal report increases the accuracy of identifying weak visual stimuli to which S had been previously aversively conditioned. These authors also found that Ss who had been instructed to delay their response showed the highest percentage of galvanic skin response (GSR) discrimination to the aversively conditioned stimulus when that stimulus was too weak to elicit a correct verbal identification. According to the authors,[3] the autonomic activation contributed additional information to aid veridical cognitive identification.

From the Menninger Foundation, Topeka, Kans.

Supported by Grants MH-8086 and MH-07962 from the National Institute of Mental Health, U. S. Public Health Service, and by funds from the John Harper Seeley Foundation.

The author gratefully acknowledges the help of a number of colleagues, especially Dr. Philip S. Holzman, Dr. Elmer Green, and Miss Lolafaye Coyne.

*Present address: Prince George's Hospital, Cheverly, Md.

Received for publication June 27, 1966.

Barry[1] studied the effects of prolonged deliberation in a group of subjects presented with the task of differentiating 7 standard weights ranging from 300 gm. to 600 gm. The Ss were required to give an immediate judgment of heaviness on a 1-to-7 scale and to consider their own statements during the 5 sec. following the presentation of the stimulus. During this period of deliberation they could change their statements if those statements seemed incorrect. The author found that in the majority of cases the changed statements were more accurate than the immediate judgments.

The studies reported here test whether allowing time for autonomic reaction to one's own voice to reach a maximum improves the accuracy of Ss' identification of their own voices. We were also interested in particular patterns of psychophysiological response to one's own voice when that voice is correctly recognized and when it is not recognized.

Experiment 1

Method

Subjects

A group of 110 Ss, 73 women and 37 men, participated in the experiment. They were paid volunteers drawn mainly from church organizations. Their ages ranged from 18 to 57 years. One hundred and one had had at least a high school education. Their experience with tape recorders varied: 92 of the Ss had heard their voices at least once, 15 had not had this experience previously, and 3 were used to hearing their own voices more than once a month.

Procedure

The procedure was as follows:

1. *S* recorded a 1- to 2-sec. sample of his voice. The phrase used was "a long, flowing beard," recorded at 60±5 db. (point of reference 0.0002 dynes × sq. cm.).

2. The voice sample, along with 19 other voices saying the same phrase, was played to the subject. *S*'s own voice was always in the twelfth position in the series.

3. Fifteen seconds after the presentation of each auditory stimulus, a buzzer sounded, signaling that *S* was to identify whether the voice he heard was his own or not.

4. Galvanic skin resistance (GSR), galvanic skin potential (GSP), and 2 channels of the electroencephalograph (EEG) were recorded.

S's task was to identify each of the 20 voices as his own or not after a delay of 15 sec. following hearing each sample.

To prevent anticipatory psychophysiological responses, the intervals between stimulus presentation were varied from 23 sec. to 78 sec. in a prearranged randomized order.

Apparatus

The subject was seated in a small room that contained a comfortable chair and 2 tables on which the microphone, speakers, and other accessory equipment were placed. A 1-way window, which separated this room from an adjacent one where the recording apparatus was placed, facilitated the experimenter's observation of the subject's behavior during the experiment.

We used 2 polygraphs: a 5-channel Offner Type "R" (Offner Division, Beckman Instruments, Inc.) to register GSR and GSP, and a 2-channel Offner Type "RS" polygraph to record the changes in EEG.

For GSP we used EEG-type silver-chloride electrodes. For the GSR the electrodes were coins of silver about the size of a quarter passing a current density of 10 μamp./sq. cm.; 2 electrodes were fastened to the right palm to allow contact with the thenar and hypothenar regions of the hand. The "active" GSP electrode was placed on the thenar region of the left hand and referred to an electrode placed on the right ear lobe. EEG recordings were made between the left ear lobe and the right parietal region (A_1-P_4) and between the 2 parietal regions (P_3-P_1), according to the 10–20 electrode system of the International Federation.

An RCA continuous-loop tape recorder (Radio Corporation of America) was used for playing back the voices. It consisted of a cartridge playback Type RT-7B and a recorder amplifier Type BA-7B.

Results

Identification of Own Voice

Table 1 shows the incidence of recognition of own voice when verbal response

TABLE 1. COMPARISON OF VOICE RECOGNITION BY ROUSEY AND HOLZMAN SS WITH SS IN PRESENT EXPERIMENT

	Recognized		*Nonrecognized*		
Group	*No.*	*%*	*No.*	*%*	*Total No.*
Rousey & Holzman[4]	63	38	101	62	164
Present study	61	55	49	45	110
TOTAL	124		150		274

Chi square: 7.044 (1 df) $p<0.01$.

was delayed 15 sec. Of the 110 Ss, 61 (42 women and 19 men), or 55%, correctly recognized their own voices and 49 (45%) did not. These figures are clearly different from those reported by Rousey and Holzman.[4] These authors requested an immediate verbal report from their Ss and found that 62% of them failed to recognize their own voices. The difference between the incidence of nonrecognition and recognition in our experiment and in theirs was evaluated by a chi-square test. A significant difference between them is apparent ($\chi^2 = 7.044$, $p<0.01$). Thus it seems likely that delaying S's response 15 sec. significantly increases the rate of own voice recognition over immediate response.

Neither age, sex, years of education, nor previous experience with tape recorders significantly affected the recognition of Ss' own voices. Chi-square tests failed to show any significant relationship between each of these factors and Ss' recognition or nonrecognition of their own voices.

Psychophysiological Response to Own Voice

Holzman *et al.*[2] found significantly greater psychophysiological activation in response to their Ss' own voices than to other voices as measured by GSR, frontalis muscle electromyogram (EMG), and finger plethysmogram. These results were highly significant when Ss recognized their own voices and showed a similar trend when Ss failed to recognize their own voices. The method for obtaining an index of recognition in that experiment was to inquire about recognition after the experiment was completed. Room for error in such a procedure is, of course, great.

In our experiment we wanted to compare Ss' psychophysiological reactions to their own and to others' voices when a delay of 15 sec. before giving the verbal report was required from the Ss. For GSR and GSP we obtained the mean amplitude of the reactions to Voices 7 to 11 and 13 to 17, and compared them with reactions to Voice 12 (S's own voice). The first 6 voices and the last 3 (18 to 20) were not included in this analysis. We excluded the reactions to the first 6 voices as reflecting unspecific orienting responses. We eliminated from consideration reactions to the last 3 voices (18 to 20) to exclude the effects of habituation to the experiment.

1. GSR. We determined each S's skin-resistance level before the presentation of any stimuli, after 5 min. of rest, as a reference for comparing continuous changes in skin resistance during the experiment. We computed 2 measures as dependent variables: (1) response time, measured in seconds, defined as the interval between the beginning of the response and the response maximum; and (2) the amplitude measured in kilohms from the prestimulus level to the maximum response expressed as a residual or deviation from the regression line. Change scores were computed as follows: (1) the correlation between pre-

stimulus values and the difference between the prestimulus and maximum response values was obtained. For GSR this value was $r = 0.11$; for GSP this value was $r = 0.22$; (2) residual scores, i.e., deviation from the regression lines representing these correlations, and thus free of the influence of prestimulus values, were then computed.

The data for GSR response time were subjected to an analysis of variance. Table 2 summarizes the significant results. There was a significant difference among voices, with Voice 12 (S's own voice) showing the longest response time (3.85 versus 2.88 sec. for S's own voice versus Voices 7 to 11 ($p<0.001$), and 3.85 versus 2.57 sec. for S's own voice versus Voices 13 to 17 ($p<0.001$). Even though there was a significant difference between S's own voice and other voices, S's own voice differed from the other voices more significantly in the group of recognizers than in the nonrecognizers.

The analysis of variance of the amplitude of GSR for S's own and other voices, as summarized in Table 2, shows that S's voice (Voice 12) produced a significantly greater reduction in skin resistance than did voices of other people, whether or not Ss recognized their own voices.

2. GSP. As with GSR, we computed 2 measures as dependent variables: (1) response time in seconds from the beginning of response to the response maximum and (2) amplitude of the response maximum, whether negative, positive, or biphasic. For the biphasic response, we measured the difference between the minimum of the negative deflection to the peak of the positive one. All raw scores are expressed in millivolts.

The data for GSP response time were subjected to an analysis of variance, and the summary is shown in Table 2. As with GSR, significant differences emerge between S's own voice and the other voices, with S's own voice showing the longest response time. A significant sex difference appears, showing that as a group women took longer to reach the response maximum.

For the analysis of variance of GSP amplitude scores, also summarized in Table 2, the raw scores were transformed into residual scores in the manner described for GSR. This transformation was undertaken as a precaution against the influence of the relationship between the prestimulus values and the raw amplitude scores. S's own voice again showed the highest response amplitude and was significantly different from Voices 7 to 11 ($p<0.001$) and from Voices 13 to 17 ($p<0.001$). These results were maintained whether Ss recognized their own voices or not. However, there was a slight tendency for the differences to be greater for recognizers as indicated by the recognition-nonrecog-

TABLE 2. SUMMARY OF SIGNIFICANT P VALUES OF VARIANCES OF Ss' PSYCHOPHYSIOLOGICAL RESPONSES TO OWN AND OTHER VOICES

	Psychophysiological response			
Variance source	*GSR response time*	*GSR amplitude*	*GSP response time*	*GSP amplitude*
Men–women (A) Recog.–nonrecog. (B)	—	<0.10	<0.025	—
Voices (C)	<0.001	<0.001	<0.001	<0.001
B × C	<0.025	—	—	<0.10

TABLE 3. ANALYSIS OF VARIANCE OF EEG ALPHA BLOCKING BEFORE AND AFTER LISTENING TO OWN AND OTHER VOICES BY MEN AND WOMEN Ss

Source	*Ss*	*df*	*F*
Uncorrected			
Men–women (A)	249.25	1	0.67
Recog.–nonrecog. (B)	193.23	1	0.52
A × B	20.39	1	0.06
Error AB	31485.29	85	
Corrected			
Voices (C)	3035.32	2	14.72*
A × C	19.59	2	0.10
B × C	271.47	2	1.32
A × B × C	342.52	2	1.66
Error C	17521.68	170	

SUMMARY OF MAIN EFFECTS

	q†
Own voice vs. 7–11	8.35‡
Own voice vs. 13–17	6.23‡
Voices 7–11 vs. 13–17	2.12

*$p<0.001$.

†$q_{r,k} = \dfrac{\chi_i - \chi_j}{\sqrt{\dfrac{s^2\ \text{error}}{n}}}$.

‡$p<.01$.

nition by voice interaction and an analysis of simple effects. In spite of the differences found between Voices 7 to 11 and 13 to 17, reflecting gradually diminished amplitude of response, probably due to habituation, S's own voice produced the greatest deflection in skin potential, interrupting that trend toward diminished response.

3. EEG Alpha Blocking. We took samples of EEG recordings 10 sec. prior to and 15 sec. after each voice presentation. For our purpose we defined alpha waves as any group of 3 or more symmetrical waves, with a frequency from 8 cps to 13 cps, and with an amplitude greater than 33% of the average of the largest group of 3 or more alpha waves found in each subject. We considered alpha waves to have been blocked when the registered waves did not fit the definition given above.

We measured the percentage of alpha blocking during the 10 sec. preceding the stimulus and during the 10 sec. after the stimulus was presented. Change scores for alpha blocking were computed in a manner similar to that described for GSR. The correlation between prestimulus and poststimulus alpha blocking was 0.26. Table 3 shows the analysis of variance of the amount of change in alpha blocking, from pre- to poststimulus, expressed as a residual. As with the autonomic data of GSR and GSP, there was significantly greater increase in alpha blocking after S's own voice than after Voices 7 to 11 ($p<0.01$) and Voices 13 to 17 ($p<0.01$). No differences were found between Voices 7 to 11 and 13 to 17. Again these differences were obtained for both Ss who did and did not recognize their own voices.

We subjected these data to a further analysis. We divided the forced choice (yes or no) responses into 4 groups: (1)

TABLE 4. T-TEST COMPARISONS OF AMPLITUDE OF PSYCHOPHYSIOLOGICAL RESPONSES ACCOMPANYING FOUR POSSIBLE VOICE RECOGNITION RESPONSES: TRUE AND FALSE POSITIVES AND TRUE AND FALSE NEGATIVES

Group	*GSP*	*GSR*	*EEG*
Correlated			
True (+) vs. False (+)	3.38*	2.51‡	1.51
True (+) vs. True (−)	5.25†	5.27†	4.26†
False (−) vs. True (−)	5.10†	5.74*	3.39*
False (−) vs. False (+)	1.12	0.53	1.22
False (+) vs. True (−)	2.59‡	3.53†	2.85§
Uncorrelated			
True (+) vs. False (−)	17.47†	1.05	0.02
	MEAN TABLE‖		
True positive	133.67	76.75	67.15
False negative	114.63	66.90	66.64
False positive	102.13	60.05	64.40
True negative	95.57	45.19	58.33

Because of the limitations of response possibilities, instead of analysis of variance a t test was used to evaluate the significance of differences. For own voice, *S* could give as an answer only a true positive or false negative response, and for "other" voices, *S* could give only true negative or false positive responses. To compare the psychophysiological responses accompanying true positive and false negative responses an uncorrelated t test was used because different *Ss* are necessarily involved. For all the other comparisons, correlated t tests were used because same *Ss* were involved.

*$p<0.005$.
†$p<0.001$.
‡$p<0.025$.
§$p<0.01$.
‖Expressed in residual score units.

true positives, in which Ss correctly recognized their own voices; (2) true negatives, in which Ss correctly reported the voice as not their own; (3) false positives, in which Ss incorrectly identified the voices as their own; and (4) false negatives, in which Ss did not recognize their own voices. The residual scores for the amplitudes of the GSR and GSP and for the change in the amount of alpha blocking before and after stimulation was used to compare these 4 groups of responses.

For the GSR and GSP, there was a significant difference between each pair of groups except for the false (−) versus false (+). The difference between true (+) versus false (−) was significant only for the GSP. These results are shown in Table 4. A hierarchy in the magnitude of the responses can be seen from the mean table. The true positive verbal reports were accompanied by the largest responses and true negatives by the smallest responses, with false negatives and false positives in that order of magnitude falling between.

Similarly, alpha blocking accompanying true negative responses was significantly smaller than that accompanying each of the other response possibilities.

Experiment 2

In Experiment 1 the improved effect on judgment of the delayed response was tested by comparing incidence of correct recognition in Rousey and Holzman's groups,[4] in which immediate judgments

were called for (38% accuracy), with that of our group, in which delayed judgment was called for (55% accuracy). We wished to know if, in the *same* Ss, accuracy of own voice recognition is adversely affected if immediate response is required.

For this study, the same Ss who participated previously were presented with the same discriminatory task of voice recognition; but instead of asking them to delay their verbal report, they were asked to identify the voices as "own voice" or "not own voice" *immediately* after hearing each recorded voice. We expected to find that a significant number of Ss who, under delayed-response conditions correctly recognized their voices, would now fail to recognize them when an immediate report was required. On the other hand, we did not expect a significant number of Ss to change from an incorrect judgment in the first experiment to a correct one in the second in spite of the possible effects of practice that could influence the Ss' recognition of their own voices and increase the probability of change from nonrecognition to recognition.

Method

Of the 110 Ss who had participated in the first experiment, 65 volunteered to return 1 year later to take part in the present experiment. The proportion of women and men, and of recognizers and nonrecognizers from the first study that participated in the present study are compared in Table 5. A higher proportion of recognizers from the first study participated in the present study.

We followed the same procedure as in our first study. Briefly, the voices of 19 people, plus that of the S, saying the same phrase, were played to the Ss. They were instructed to report immediately after hearing each voice whether the voice they heard was their own or that of another person.

Results

The results are shown in Table 6. The number of Ss who recognized their own voices in the second study (32) was smaller than the number of subjects who

TABLE 5. PROPORTION OF MEN AND WOMEN; RECOGNIZERS AND NONRECOGNIZERS WHO PARTICIPATED IN THE TWO STUDIES

	Recognizers		*Nonrecognizers*		*No.*		*%*	
	Study 1	*Study 2*	*Study 1*	*Study 2*	*Study 1*	*Study 2*	*Study 1*	*Study 2*
Women	42	27	31	17	73	44	66	68
Men	19	14	18	7	37	21	34	32
No.	61	41	49	24	110	65	—	—
%	55	63	45	37	—	—	—	—

TABLE 6. ANALYSIS OF THE CHANGE IN CORRECT RECOGNITION FROM THE FIRST STUDY TO THE SECOND ONE

		Second study		
		Nonrecognizers	*Recognizers*	*No.*
First study	Nonrecognizers	16	25	41
	Recognizers	17	7	24
	No.	33	32	65

Chi square (1 df) 3.52 $p<0.05$ (1-tailed test).

recognized their own voices in our first study (41). A chi-square test for change, the McNemar Test,[5] was used to evaluate the change in the verbal reports of individual Ss. Significantly more Ss changed their judgments from correct under delayed response to incorrect under immediate response than changed from incorrect to correct ($p<0.05$, 1-tailed test). Of the 41 recognizers in the first study, 16, or 39%, became nonrecognizers when immediate response was required. On the other hand, of the 24 nonrecognizers in the first study, only 7, or 29%, became recognizers in the second study.

Conclusions

1. The results clearly show that when Ss with cultural characteristics similar to those of the group tested are confronted with a discrimination task—in this instance identifying one's own recorded voice—delaying verbal response to the discrimination significantly increases accuracy of judgment. The results are, thus, consistent with those of Reiser and Block[3] for discrimination of aversively conditioned "neutral" stimuli. The results of these studies confirm our hypothesis that when Ss are confronted with the discriminatory task of identifying their own recorded voices, delaying the verbal response increases significantly the accuracy of the Ss' identification of their own voices. Apparently the effects of experience gained by the Ss in listening to their own voices by participating twice during the same year in similar experiments did not significantly affect the probability of recognizing their own voices during the replication of the experiment.

2. Psychophysiological responses to own voices were significantly greater than to others' voices, whether or not Ss were consciously aware of listening to their own voice. These responses included GSR, GSP, and EEG alpha blocking. These results afford clear confirmation for those of Holzman *et al.*[2]

The findings suggest that autonomic physiological responses and changes in cortical activity discriminate significant verbal stimuli independent of S's conscious judgments. There seem to be two kinds of responses closely related, one a response to the stimulus and the other to the S's awareness of the nature of the stimulus. The perceptual experience is probably the result of a complicated interaction of these two responses, involving reciprocal feedback of autonomic and cognitive information. It seems highly likely that the autonomic feedback from the activation produced by own voice is recruited to adaptive use in aiding recognition when time is allowed for such feedback to become effective.

Prince George's Hospital
Cheverly, Md.

References

1. Barry, H. Effects of prolonged deliberation on psychophysical judgment. *Amer J Psychol* 77:270–275, 1964.
2. Holzman, P. S., Rousey, C., and Snyder, C. On listening to one's own voice: Effects on psychophysiological responses and free associations. *J Personality Soc Psychol* 4:431, 1966.
3. Reiser, M. H., and Block, J. D. Discrimination and recognition of weak stimuli. III. Further experiments on interaction and autonomic feedback mechanisms. *Psychosom Med* 27:274–285, 1965.
4. Rousey, C., and Holzman, P. S. Recognition of one's own voice. *J Personality Soc Psychol* In press.
5. Siegel, S. *Nonparametric statistics for the behavioral sciences.* McGraw-Hill, New York, 1956, p. 63.