

3. POWERS, H. E. An application of the Marston introversion-extroversion scale. *J. Educ. Psychol.*, 1928, **19**, 168-174.

*Columbia University*  
*New York City*

SOCIAL AGREEMENT ON PERSONALITY TRAITS AS  
JUDGED FROM SPEECH<sup>1</sup>

HAROLD C. TAYLOR

A. THE PROBLEM

It should be obvious that the speech of another person—along with his dress, mannerisms, physical build, and physiognomy—greatly influences our judgments concerning his personality. The mere “hello,” “thank you,” and “good-bye” of a telephone operator often give us quite complex notions of her personality. We speak of a radio announcer as cultured, egotistical, nervous, or whatnot, entirely apart from the content of his utterances. Psychologically trained persons are apt to take little stock in the validity of such judgments, but they are real, nevertheless. And just as studies of judgments from photographs have revealed the different degrees of social agreement in such judgments on various traits, as well as the general inaccuracy of all such judgments, a similar study might be expected to show the situation with regard to speech.

Pear's (1) study of voices as heard on the radio utilized a great many listeners, but few speakers, and was confined to judgments of a few traits, such as age, sex, occupation, and the section of England in which the speaker was reared. No other experiments of this nature have been reported in the literature.

This study confined itself to these three questions: (1) To what extent do people agree in ascribing a given trait to another, with speech as their only criterion? (2) How accurate are such judgments? (3) Are the traits on which more people agree also the ones which are judged more accurately?

B. PROCEDURE

Twenty young men acted as subjects. Eight of these were from the least neurotic fifty (as judged by the Thurstone Personality Schedule)

---

<sup>1</sup>The writer is deeply indebted to Professor Edward Sapir, under whose auspices this work was done at the University of Chicago, for his personal encouragement and for his aid in developing the problem. For the statistical procedures employed and for technical advice, the writer is indebted to Professor L. L. Thurstone; and to Mr. S. N. Treviño of the University of Chicago for making the disk records used in this study.

in the 1930-31 freshman class at the University of Chicago. Seven were from the most neurotic fifty, and five were upper classmen chosen at random. Each subject read a three-minute newspaper editorial on preparedness, and his reading was recorded on a Speakphone aluminum disk.

A questionnaire was prepared, consisting of 136 questions, statements, or words descriptive of personality traits. Thirty-two of these were from Freyd's list of introvert characteristics and 25 were from the Thurstone Personality Schedule. The remaining items were words commonly used to describe people. Each subject filled out this questionnaire, indicating by a "plus" sign every trait he thought he possessed, by a "minus" sign every trait he thought he did not possess, and by a question mark those on which he could not decide.

Each of the twenty records was then played to at least twenty auditors, who, after hearing a given record, filled out the same questionnaire, rating the subject whose voice they had just heard. Auditors were not permitted to use the question mark. Four hundred fifty-five questionnaires were secured from auditors.

### C. RESULTS

1. The first problem was to ascertain the degree of social agreement shown in judging personality traits, regardless of the accuracy of such judgments. The measure of the degree of social agreement was derived as follows, using item 1—"effeminate"—as an example: 35% of all the 455 auditors marked that item "plus." Taking 35%, therefore, as the mean chance expectancy for "plus" marks on this trait, the formula

$\sqrt{\frac{pq}{n}}$ , or  $\sqrt{\frac{.35 \times .65}{20}}$  gives the standard deviation or "spread" from

35% of "plus" marks which chance would account for with twenty cases. If the actual standard deviation of these twenty cases whose mean is 35% is appreciably greater than that which chance will account for, then to that extent it may be assumed that the auditors were making a real discrimination. It would mean that, for a given trait, some subjects received a great many "plus" ratings, other subjects very few, and that these groupings were on the whole too marked to have arisen by chance. For item 1, then, the spread to be expected by chance is 10.66%, the actual spread is 27.46%, and the difference is thus 16.80%. To go one step further, this difference is 5.41 times its probable error; and at a superficial glance "effeminate" seems therefore established statistically as a trait on which there is a high degree of social agreement, when it is judged from speech. However, because of the small  $n$  of 20, any single

one of these critical ratios may be highly inaccurate, and the data seem hardly to justify any comparisons of the individual items of the questionnaire.

Table 1 presents a summary of the critical ratios for the 136 items, each computed as described above. These critical ratios represent the degree of certitude that the auditors made a real discrimination, not attributable to chance, between subjects with respect to a given trait. Thus, for example, the apparent degree of social agreement found in judging the best 72 items would be as low as the mean chance expectation less than once in a thousand times. It is clear that there is a marked degree of social agreement on many personality traits among judges whose only criterion is the subject's speech.

2. The second problem concerns the accuracy of these judgments based on speech, using agreement with subjects' self-ratings as the criterion. Three different measures of accuracy were employed.

*a.* It will be recalled that there was included in the questionnaire a short test of neurotic tendency, consisting of 25 items from the Thurstone Personality Schedule. Each of the twenty subjects has, therefore, two scores on this test: one from his own self-ratings, the other from the pooled ratings on these items of the twenty or more people who heard the record of his voice. The correlation between auditors' ratings and subjects' self-ratings on this test is  $+0.06$ . That this low relationship is not due to unreliabilities in the test is evident, since the reliability coefficient of the test is 0.99 when used by auditors, 0.94 when used by the subjects themselves. Clearly, there is no apparent relationship between a subject's own neurotic score and the score ascribed to him by people who judge him from his speech alone.

*b.* There was also a short introversion schedule of 32 items included in the questionnaire. The correlation between subjects' own scores and the auditors' ratings on this test is  $+0.03$ , indicating no relationship whatever. While there is definite social agreement that some voices ought

TABLE 1  
EXTENT TO WHICH AUDITORS' DISCRIMINATION EXCEEDS THE MEAN CHANCE  
EXPECTED DISPERSION OF JUDGMENTS

	Number of items	Critical ratios
Highest	23	4.50 or over
"	55	3.75
"	72	3.00
"	88	2.25
"	111	1.50
"	121	1.00

to belong to introverts, such voices appear to be possessed as often as not by extroverts.

c. As a third measure of agreement between auditors' judgments and subjects' self-ratings, tetrachoric  $r$ 's were computed for 125 of the 136 traits. Insufficient data were available for calculating the other 11. The cells of the fourfold table were filled as follows: on the trait "effeminate," 35% of all auditors recorded "plus." Then if more than 35% of the auditors of Subject I marked him "plus," he was assigned a "plus" in the auditors' ratings; and if he marked himself "plus," he was recorded in the "plus-plus" cell of the fourfold table. These 125 coefficients ranged from  $+0.84$  to  $-0.82$ . The mean was  $-0.07$ , and the standard deviation 0.39.

From the mean of  $-0.07$  it is clearly evident that there was no general tendency for the auditors' opinion to agree with the subjects' opinion of themselves. The further question arises as to whether the high correlations have any significance. While the probable error of a number of these indicates them to be statistically reliable, the P.E. is itself so unstable with an  $n$  of twenty that 125 tetrachoric  $r$ 's may be expected by chance to vary almost as much as these did if the true correlation is zero.

However, it may be that these correlations are to some extent real and that traits are themselves widely distributed in the extent to which they are revealed by speech. That is, most traits may be very slightly revealed, some may be strongly revealed, while in still others the apparent significance of the speech cues may be quite misleading. But it must apparently be concluded from these data that, in general, when those who hear a subject's speech agree in ascribing to him a given trait, he himself is as likely as not to deny its possession. This clearly does not mean that in real life we cannot tell anything about a person by listening to him talk. It suggests that while under such conditions we do base personality judgments upon speech and voice, whatever accuracy such judgments may possess is contingent upon what the person says rather than upon the general characteristics of his speech and voice.

3. The third problem was whether the traits on which people agree best among themselves are also the ones on which they agree most with the subjects' own self-ratings. An "index of agreement" was devised to show to what extent the auditors agreed among themselves in their assignments of each trait, and these indices were correlated with the tetrachoric  $r$ 's. The coefficient thus obtained with  $-0.26$ , which is 4.5 times its probable error. There is thus a slight but statistically significant tendency for the auditors to be most consistent in their judgments when they agree least with the subjects themselves.

## SUMMARY AND CONCLUSIONS

1. There is clearly a high degree of social agreement in judging the personality traits of people with speech as the only guide.
2. Social judgments thus based on speech bear no relationship to the judgments of the subjects themselves so far as the present data are concerned. This is true of judgments of neurotic tendency and of introversion and of the mean accuracy of judgment of 125 diverse traits. The few cases of strong agreement or disagreement with subjects' self-ratings are probably not of great significance.
3. There is a tendency for the auditors to be most consistent in their judgments when they agree least with the subjects' self-ratings. It is at least clear that their degree of consistency in judging a given trait bears no positive relationship to their accuracy in judging that trait.

## REFERENCE

1. PEAR, T. H. *Voice and personality*. London: Chapman & Hall, 1931. Pp. xii+247.  
*Yale University*  
*New Haven, Connecticut*

## DISCRIMINATION: A STUDY OF SOCIAL DETERMINANTS

HOWARD WOOLSTON

In a previous paper, the writer (1) showed that social conditions affect the exclusiveness of judgments concerning peculiar persons in certain relations. Therein no attempt was made to differentiate the value of these factors. Using the same data, we shall now try to indicate how the weight of such predisposing circumstances can be measured.<sup>1</sup>

The original inquiry sought to find how students regard the admission of ten atypical personalities into ten different associations. If they admitted everybody to all these relations, their score was one hundred inclusions. If they rejected everybody, their rating was a hundred exclusions. Summing up a hundred such responses, we found nearly 6600 acceptances and more than 3400 rejections. In other words, the exclusive judgments of the group comprised about 34% of the total number of expressions. We, therefore, considered this average as the best representative figure to show how student opinion responded to the idea of

---

<sup>1</sup>Seven of the original hundred data sheets were discarded, because of incomplete or doubtful entries. This elimination alters some of our previous results, but does not affect the main conclusions.