The tone of voice in which therapists spoke about their alcoholic and/or drug-abusing patients was employed to predict the tone of voice in which these same therapists would speak to these same patients. Clips of speech ranging from 10 s to 1 min were content filtered and rated by judges. The results of zero-order correlational analysis, multiple regression, and canonical correlational analysis all showed clearly that predictions of therapists' tone of voice could be made with discriminant as well as predictive validity. Predictions were not only very significant statistically, but of practically meaningful magnitudes. Therapists who spoke about patients in a coldly autocratic way tended to speak to these patients in a coldly professional way. Therapists who spoke about patients in a warm and caring way tended to speak to these patients in a warm and honest tone of voice.

One such study showed that hostility in a referring clinician's tone of voice in speaking about alcoholic patients served as a significant postdictor of referral success. The more hostile the referring physician, as judged from brief (1.5 min) content-filtered speech samples, the less successful was that physician in getting alcoholic patients to go to an alcoholism treatment facility (Milmoe, Rosenthal, Blane, Chafetz, & Wolf, 1967). It should be noted that these content-filtered speech samples were of clinicians speaking about patients, not directly to them. Our interpretation of the results of this research, however, suggested that clinicians probably spoke to patients in a hostile, cold way if they spoke about those same patients in a hostile, cold way. That seemed to be a reasonable inference but one on which there were no available data.

The purpose of the present study was to investigate the degree to which clinicians' tone of voice in speaking about their patients could be used to predict clinicians' tone of voice in speaking to those same patients. If tone of voice in speaking about patients showed such predictive validity, it would do much more than simply support the interpretation we had offered some 15 years earlier.

Indeed, even moderate predictive correlations could be of great value to clinical researchers, supervisors, teachers, and faculty serving on admissions committees. Because
there may be serious logistic and ethical problems associated with studying interactions of clinicians (and student clinicians) with their patients, it would be extremely valuable to learn whether the way clinicians spoke about their patients reflected the way they spoke to their patients. If it were established that the tone of voice in which therapists spoke to patients affected patients' outcomes, researchers could, at least sometimes, use therapists' tone of voice in speaking about patients as a simulation of therapists' tone of voice in speaking to patients. This would prove to be a valuable additional avenue to the study of clinical processes that would help address the logistical and ethical problems sometimes encountered in the study of clinician–patient interactions. In a similar vein, future research might suggest the value to clinical supervisors and to clinical faculty admissions committees of assessing the tone of voice in which student clinicians, supervisees, or applicants for training speak about their patients. Clinical experience suggests that such assessments are already made, though usually only implicitly.

Method

Therapists and Patients

Twenty-one (10 male and 11 female) psychologists, psychiatrists, counselors, and social workers (both inpatient and outpatient staff) spoke both to and about 98 patients who had participated either in an inpatient or an outpatient 5-week program for alcohol (and drug) abusers. The 56 male and 42 female patients consisted of 46 inpatients and 52 outpatients. Although all patients were substance abusers (primarily alcoholics), the subject sample was heterogeneous with respect to sex, educational background, and socioeconomic status. In addition, there was considerable heterogeneity among the subjects in terms of character structure and diagnosis accompanying the alcoholism, covering the full range from neurotic to psychotic and encompassing a wide diversity of DSM-III (Diagnostic and Statistical Manual of Mental Disorders, 3rd ed., American Psychiatric Association, 1980) categories (including manic depression, schizophrenia, and obsessive compulsive, generalized anxiety, adjustment, and dysthymic disorders). Further details of therapist and patient selection are described elsewhere (Blanch, Rosenthal, & Vannicelli, 1983).

Interview Contexts

Typically, each therapist spoke to both female and male patients (either inpatients or outpatients) for whom they held varying levels of expectation for improvement. Patients were told that this was a special study of the alcohol treatment program at the hospital. As part of the study, patients were told that staff were interested in getting impressions of the program from those people who were currently participating in the program and from those who had participated in the past. Patients were asked to talk to a staff member for 15 min about their impressions of the following: (a) the setting (buildings, grounds, etc.), (b) the scheduling of activities, (c) the therapy groups, (d) the involvement of family members, (e) the hall meetings, and (f) whether or not they would recommend the program to a relative or friend.

Each therapist also spoke about each of these same patients to a member of our research staff. They were asked to talk about their patients for approximately 3–5 min. Therapists were instructed to discuss anything that came to mind about each of the patients or about their experiences with particular patients that led them to feel a certain way about them. The actual sequence of the therapist's discussions of the patients (high—low expectations, male—female) was determined by a prearranged randomization schedule that was different for each therapist.

Interview Recordings

The 98 interviews with patients were audiotape recorded and ranged in length from approximately 4 min to 37 min, with a mean length of approximately 17 min. For purposes of the present study, three different tapes were derived from the original recording employing the following method: From each interview, the first 10 s of the therapist's speech was assembled and rerecorded onto one tape; 10 s of the therapist's midmost speech was assembled and rerecorded onto a second tape; finally, the last 10 s of the therapist's speech was assembled and rerecorded onto a third tape. (We decided on 10-s clips because these generally captured the bulk of therapists' comments to patients.) The order of the 98 clips on each of the three tapes was randomized. Because we were interested only in therapist's tone of voice in this study, the patient's speech was excluded from the derived tapes.

The final result was three tapes containing ninety-eight 10-s segments or clips of speech from the beginning, middle, and ending portions of each interview. These master tapes were then content filtered, a process that removes from the tape the high frequencies on which word recognition depends but that preserves sequence and rhythm (Rogers, Scherer, & Rosenthal, 1971).

The 98 interviews about patients were also audiotape-recorded and ranged in length from about 3 to 5 min, with a mean length of approximately 3 min. Three master tapes were again derived from the original recordings on the following basis: From each interview, the first 20 s of the therapist's speech was assembled and rerecorded onto one tape; 20 s of the therapist's midmost speech was assembled and rerecorded onto a second tape; and finally, the last 20 s of the therapist's speech was assembled and rerecorded onto a third tape. (We decided on 20-s clips because these seemed better to capture the bulk of therapists' comments about patients.) Again the order of the clips on each of the three tapes was randomized.

The final result was three tapes containing ninety-eight 20-s segments or clips of speech from the beginning, middle, and ending portions of each interview. These master tapes were also content filtered.
Judges and Ratings

Twelve (6 male and 6 female) undergraduates at Harvard University were each paid to rate all three tapes (beginning, middle, and ending segments) of therapists talking to patients, and another 12 undergraduates (6 male and 6 female) were each paid to rate all three tapes of therapists talking about patients. Judges were randomly assigned to one of three counterbalanced conditions representing the order in which they would rate the three tapes (i.e., ABC, BCA, CAB). Judges were told that they would hear therapists talking and that they would be able to understand only the speaker's tone of voice. Consistent with our earlier experience with this type of study, judges were given no special training for their task (Rosenthal, 1982). All judges rated all the segments (98 beginning, 98 middle, and 98 ending segments) on 10 dimensions—not warm—warm, not hostile—hostile, not anxious—anxious, not dominant—dominant, not empathic—empathic, not competent—competent, not optimistic—optimistic, not professional—professional, not honest—honest, and not liking—liking. Each segment was played once, and all judges were given 15 s to make the 10 ratings. Each rating was made on a scale running from 1 (not at all warm) to 9 (very warm). Because ratings of the beginning, middle, and ending segments were substantially correlated (median \( r = .49 \) for the shortest clips), these ratings were combined prior to subsequent analyses (see Blanck, Rosenthal, Vannicelli, & Lee, 1983, for a detailed discussion of the various types of phase-to-phase and judge-to-judge reliabilities).

Principal Components Analysis

The mean of all judges’ ratings of the tone of voice in which therapists spoke to and spoke about their patients were intercorrelated separately, and a principal components analysis was computed for each of these 10 \( \times \) 10 correlation matrices of mean judges’ ratings. For the ratings of therapists’ tone of voice in speaking to their patients, this analysis yielded four interpretable factors after varimax rotation: (a) professional—competent (consisting of professional and competent, optimistic, and professional), (b) warmth (consisting of warm, not hostile, not dominant, empathic, and liking), (c) anxious, and (d) honest. Each new factor-based variable was defined as the mean rating of the variables included in that factor with the sign of the loading taken into account. For example, the warmth variable was defined as the mean rating of warm and liking. Consistent with our earlier experience with this type of study, judges were given no special training for their task (Rosenthal, 1982). All judges rated all the segments (98 beginning, 98 middle, and 98 ending segments) on 10 dimensions—not warm—warm, not hostile—hostile, not anxious—anxious, not dominant—dominant, not empathic—empathic, not competent—competent, not optimistic—optimistic, not professional—professional, not honest—honest, and not liking—liking. Each segment was played once, and all judges were given 15 s to make the 10 ratings. Each rating was made on a scale running from 1 (not at all warm) to 9 (very warm). Because ratings of the beginning, middle, and ending segments were substantially correlated (median \( r = .49 \) for the shortest clips), these ratings were combined prior to subsequent analyses (see Blanck, Rosenthal, Vannicelli, & Lee, 1983, for a detailed discussion of the various types of phase-to-phase and judge-to-judge reliabilities).

Reliability of Judges’ Ratings

The reliability of the judges’ ratings of the tone of voice variables was computed by means of intraclass correlations (Rosenthal, 1982). The effective reliability of the mean of the 12 judges’ ratings for the 30-s clips (three clips of 10 s each) of therapists speaking to patients ranged from .22 to .82, with a median \( r \) of .59. The effective reliability of the mean of the 12 judges’ ratings for the 60-s clips (three clips of 20-s each) of therapists speaking about patients ranged from .47 to .83, with a median \( r \) of .64. In evaluating these median reliabilities of .59 and .64, respectively, one should keep in mind that they are based on ratings of content-filtered speech of only 30-s and 60-s duration, respectively. For our present purposes the obtained reliabilities are more than adequate; however, had we increased our segment lengths to 10 min, we could have increased our median reliabilities to .97 and .95, respectively (Rosenthal, 1982).

Results and Discussion

We employed three types of analyses to address the question of whether talking about patients was predictive of talking to patients: (a) simple correlations, (b) multiple regressions, and (c) canonical correlation. We describe each analysis in turn.

Simple Correlations

For our first analysis we correlated each of the 10 variables that had been rated by judges while therapists talked to patients with each of the corresponding 10 variables that had been rated by judges while therapists talked about patients. The first column of correlations of Table 1 shows the results of this analysis. All of the correlations were positive, most were significant, and half were significant at \( p < .01 \). The median effect size \( (r) \) of .236 was also significant at \( p < .01 \) and was equivalent in practical magnitude to increasing the accuracy of prediction from 38% to 62% by means of the binomial effect size display (BESD; Rosenthal & Rubin, 1982). These correlations provide strong evidence for the hypothesis that therapists’ tone of voice in talking about patients can be used to predict significantly, and
Table 1

Predicting Therapists’ Tone of Voice While Talking to 98 Patients From Therapists’ Tone of Voice While Talking About Those Same 98 Patients: Simple Correlations

<table>
<thead>
<tr>
<th>Tone of voice variable</th>
<th>Correlation with predictor variable</th>
<th>Mdn correlation with nonpredictor variables</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warmth</td>
<td>.311***</td>
<td>.127</td>
<td>.184</td>
</tr>
<tr>
<td>Hostility</td>
<td>.094</td>
<td>.079</td>
<td>.015</td>
</tr>
<tr>
<td>Anxiety</td>
<td>.083</td>
<td>.066</td>
<td>.017</td>
</tr>
<tr>
<td>Dominance</td>
<td>.270***</td>
<td>.008</td>
<td>.262</td>
</tr>
<tr>
<td>Empathy</td>
<td>.225*</td>
<td>.052</td>
<td>.173</td>
</tr>
<tr>
<td>Competence</td>
<td>.253**</td>
<td>.045</td>
<td>.208</td>
</tr>
<tr>
<td>Optimism</td>
<td>.276***</td>
<td>.137</td>
<td>.139</td>
</tr>
<tr>
<td>Professionalism</td>
<td>.248**</td>
<td>.003</td>
<td>.245</td>
</tr>
<tr>
<td>Honesty</td>
<td>.048</td>
<td>.175</td>
<td>-.127</td>
</tr>
<tr>
<td>Liking</td>
<td>.154</td>
<td>.020</td>
<td>.134</td>
</tr>
<tr>
<td>Mdn</td>
<td>.236**</td>
<td>.059</td>
<td>a</td>
</tr>
</tbody>
</table>

* The difference between the medians is .177; the median of the 10 differences is .156.
* * p < .025, one-tailed. ** p < .01, one-tailed. *** p < .005, one-tailed.

with practical benefit, the tone of voice therapists are likely to employ in talking to patients. This is all the more remarkable when we recall that the total length of the speech sample of talking to patients was only 30 s and that the total length of the speech sample of talking about patients was only 60 s. Furthermore, the speech samples of talking to patients were rated by a different group of 12 judges than were the speech samples of talking about patients.

In addition to the predictive validities (reported in the first column of correlations of Table 1), we wanted also to investigate the differential or discriminant validities of our predictor variables. For example, suppose that warmth while talking to a patient correlated .311 with warmth while talking about a patient (which it did) but that warmth while talking to a patient also correlated .311 with hostility, anxiety, dominance, and so forth, while talking about the patient. The predictive validity of warmth while talking about a patient would not thereby be impugned, but the differential validity or discriminant validity would be impugned. Accordingly, for each criterion variable in turn, we computed correlations with the nine nonpredictor variables that should not have predicted the criterion variable very well. For example, for the criterion variable of warmth (while talking to patients), we computed its correlations with hostility, anxiety, dominance, and so forth (while talking about patients). If there were differential or discriminant validity, these correlations should, on the average, be lower than the specifically predictive correlations, for example, warmth while speaking about with warmth while speaking to particular patients.

The second column of correlations of Table 1 shows the median correlation of these “nonpredictive” variables with the criterion variable. None of these median correlations was significant; their median value was only .059, and only one of these, that for the criterion variable of honesty (while talking to) was larger than the corresponding predictive correlation. This finding, that it is the particular nonverbal variable judged from tone of voice while speaking about patients that best predicts the same nonverbal variable judged from tone of voice while speaking to patients, adds substantial construct validity to the predictive validity we have established.

Multiple Regressions

After establishing the basic predictive and discriminant validity of nonverbal behaviors while talking about patients as predictors of nonverbal behaviors while talking to patients, we wanted to learn how much better we might do employing multiple regression rather than just simple correlation. However, we did not want to employ 10 predictors to predict 10 dependent variables in turn. Accordingly, we employed as our dependent variables the four supervariables that had been based on our principal components analysis of the “to” segments. The supervariable of professional competence was made up of the mean of the variables of competence, optimism, and professionalism. The supervariable of warmth was made up of the mean of the variables described as warm, not hostile, not dominant, emphatic, and liking. The remaining two variables (anxiety and honesty) stood alone.

We employed as our predictor variables the seven supervariables that had been based on our principal components analysis of the “about” segments. The supervariable of professional competence was made up of the
mean of those two variables (professionalism, competence). The supervariable of warmth was made up of the mean of the variables warmth, empathy, and liking. The remaining five variables (anxiety, dominance, optimism, honesty, and absence of hostility) stood alone. These seven variables were employed as the predictor battery for each of the four dependent variables, in turn.

Table 2 shows the results of the regression analysis. Two variables significantly predicted professional competence in therapists' tone of voice in talking to their patients: talking about them in a dominant and an optimistic tone of voice. The multiple $R$ was not only significant statistically ($p = .002$) but also substantial in magnitude (.356; see Rosenthal & Rubin, 1982).

The supervariable warmth (while talking to patients) was significantly predicted by the variables of speaking about the patients in a tone of voice that was not dominant, but anxious, and warm. The last of these predictors, warmth, was a supervariable comprised of the variables warmth, empathy, and liking. The multiple $R$ of .425 was significant at $p < .001$.

Anxiety in the therapist's tone of voice in talking to patients was not significantly predicted by any battery of predictor variables. Table 2 shows that honesty in talking about a patient was significantly related to anxiety while talking to patients only when hostility was also entered into the regression equation. The trend, though not significant, was for therapists' anxiety in talking to patients to be predicted from their talking about patients in an honest but hostile manner.

Honesty in the therapist's tone of voice in talking to patients was significantly predicted by a battery of three predictors: talking about their patients in a way that was warm, not honest, and anxious. The multiple $R$ of .404 was significant at $p < .001$.

The results of the four multiple regressions showed that three of the four dependent supervariables describing therapists' tone of voice

### Table 2
Predicting Therapists' Tone of Voice While Talking to 98 Patients From Therapists' Tone of Voice While Talking About Those Same 98 Patients: Multiple Regression

<table>
<thead>
<tr>
<th>Predictor variable</th>
<th>Unique variance</th>
<th>$t$</th>
<th>$p$</th>
<th>Effect size ($r$)</th>
<th>$R$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicting professional competence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominance</td>
<td>.074</td>
<td>2.85</td>
<td>.006</td>
<td>.281</td>
<td>.356</td>
<td>6.89**</td>
</tr>
<tr>
<td>Optimism</td>
<td>.036</td>
<td>1.99</td>
<td>.050</td>
<td>.200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting warmth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominance</td>
<td>.098</td>
<td>-3.35</td>
<td>.002</td>
<td>-.327</td>
<td>.425</td>
<td>6.92***</td>
</tr>
<tr>
<td>Anxiety</td>
<td>.091</td>
<td>3.23</td>
<td>.002</td>
<td>.316</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warmth$^d$</td>
<td>.042</td>
<td>2.19</td>
<td>.032</td>
<td>.220</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honesty</td>
<td>.041</td>
<td>2.01</td>
<td>.048</td>
<td>.202</td>
<td>.211</td>
<td>2.20*</td>
</tr>
<tr>
<td>Absence of hostility</td>
<td>.026</td>
<td>-1.60</td>
<td>.113</td>
<td>-.162</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting honesty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.404</td>
<td>6.13***</td>
</tr>
<tr>
<td>Warmth$^d$</td>
<td>.125</td>
<td>3.74</td>
<td>.001</td>
<td>.360</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honesty</td>
<td>.045</td>
<td>-2.26</td>
<td>.027</td>
<td>-.227</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>.042</td>
<td>2.16</td>
<td>.034</td>
<td>.217</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ For predicting professional competence $df = 95$; for predicting warmth $df = 94$; for predicting anxiety $df = 95$; for predicting honesty $df = 94$.

$^b$ Computed as $\sqrt{\frac{t^2}{t^2 + df}}$.

$^c$ For predicting professional competence $dfs = 2, 95$; for predicting warmth $dfs = 3, 94$; for predicting anxiety $dfs = 2, 95$; for predicting honesty $dfs = 3, 94$.

$^d$ Comprised of warmth, empathy, and liking.

in talking to patients could be predicted significantly from knowing how therapists talked about patients as measured on five predictor variables—warmth, anxiety, dominance, optimism, and honesty. The remaining two predictor variables of hostility and professional competence in talking about patients did not contribute significantly to the prediction of the four factor-based dependent variables.

Figure 1 summarizes the four regression equations graphically. The four dependent variables are shown in boxes, and the five significant predictors are shown on the perimeter of the diagram. This form of display of a set of multiple regression equations has the advantage that it not only shows for each dependent variable its successful predictors but also shows for each predictor the various dependent variables to which it makes a predictive contribution.

The single substantive conclusion that seems clearest is that therapists who speak with warmth and anxiety (interpreted in this context as concern and caring) and not in a dominant manner about their patients, treat their patients with warmth and honesty as judged from their tone of voice in speaking to their patients.

Additionally, more tentative conclusions from Figure 1 are that (a) therapists speaking of patients in a dominant but optimistic manner are more likely to talk to those patients in a professionally competent manner; (b) honesty in talking about a patient may be associated with anxiety in talking to that patient; and (c) honesty after correction for any component of warmth (see Table 2) may predict dishonesty in tone of voice while talking to patients. Honesty corrected for warmth may be akin to brutal frankness, a kind of "honesty" that may not be of therapeutic benefit.

Canonical Correlation

In our final analysis we examined the overall relationship between the seven predictor variables and the four criterion variables we had earlier examined in a series of four multiple regression equations. The resulting canonical correlation $R_c$ was $0.584$, $p = 0.00022$, showing that the information in the predictor set significantly predicted the information in the criterion set. Of the four canonical correlations that could be computed, only the first was significant.

Table 3 shows the loadings on the first canonical variate for the predictor and the criterion variables. The pattern of loadings on the predictor variables suggest that therapists scoring high on this canonical variate manifest a tone of voice in talking about patients that is cold, uncaring (unanxious), perhaps overly professional, and dominating—a pattern that appears "coldly autocratic." These coldly au-
tocratic speakers about patients when speaking to patients are characterized as being cold, dishonest, and again, perhaps overly professional or "coldly professional." In short, therapists who speak of patients in a coldly autocratic way tend to speak to those patients in a coldly professional way.

The magnitudes of relationship found in the present clinical context are quite comparable to those found in other clinical contexts such as psychotherapy outcome research (Smith & Glass, 1977; Smith, Glass, & Miller, 1980) and in other areas assessing the effects of nonverbal processes of social influence such as the effects of interpersonal expectations (Rosenthal, 1966; Rosenthal & Jacobson, 1968; Rosenthal & Rubin, 1978).

Conclusion

The present study shows that therapists' tone of voice in talking about patients can be used to make significantly accurate predictions about therapists' tone of voice in talking to patients. Our research shows, furthermore, that these predictions are characterized by discriminant validity as well as predictive validity. The predictions about therapists' tone of voice are not only significant statistically but also practically meaningful in magnitude. Correlations based on the median simple, the median multiple, and the cumulative canonical correlations were .24, .38, and .39, respectively. These correlations are associated with improvements in prediction from 38% to 62%, from 31% to 69%, and from 30.5% to 69.5%, respectively (Rosenthal & Rubin, 1982). In short, these correlations, which are about as high as are usually obtained in clinical research of this type (Cohen, 1977, page 81), might well be of practical utility if their magnitude holds up on replication.

We are intrigued by the possibility that such brief clips of content-filtered speech might be useful not only for further research on psychotherapeutic process but also, in the future, for the selection and supervision of therapists as well.

References


Received December 12, 1983
Revision received February 15, 1984