

A Note on Nonprofessional Judgments of Mental Health

David A. Summers, Ph.D.*

Ted Faucher, M.A.

Sue B. Chapman, M.A.

ABSTRACT: *Ninety Ss who had either not received or not completed formal diagnostic training made mental health judgments about 60 hypothetical persons, each of whom was represented on a 7-cue behavior profile. A comparison of these judgments with those made by 24 clinical psychologists and psychiatrists failed to reveal a significant difference between groups in terms of scope of cue utilization, magnitude of judgments, or confidence. Differences between groups were obtained, however, in terms of judgment reliability and profile cue utilization. Notable in regard to the latter was the finding that nonprofessional Ss usually ignored information of a "positive" nature ("enthusiasm"), and tended to weight heavily information about violations of legal norms.*

During the past two decades, a large number of empirical studies of clinical judgment have been reported, many of which have been primarily concerned with judgment *processes*, rather than *accuracy* (Goldberg, 1968). Despite the attention that has been given to clinical judgment, it is unfortunate to note that with only a few exceptions (e.g., Hunt, Jones, & Hunt, 1957; Marabian & Reed, 1969; Yarrow, Schwartz, Murphy, & Deasy, 1955) these studies have dealt almost exclusively with judges who have received formal diagnostic training—for example, clinical psychologists and psychiatrists.

Although this emphasis on professionally trained judges has been valuable, there is evidence (Ellsworth, 1968) that certain diagnostic or therapeutic roles in a variety of mental health settings are frequently performed by "nonprofessionals" or "paraprofessionals"—that is, by individuals who have received little, if any, formal diagnostic training. In view of current and projected manpower needs in this area, it is likely that nonprofessional involvement in mental health settings will increase (Ewalt, 1967; Matarazzo, 1971). As yet, however, very little is known about how such "nonprofessionals" make judgments about the mental health of others, or about the extent to which these judgments may differ from those made by professionally trained diagnosticians. It is toward this problem that the present study is addressed.

The range of nonprofessionals who conceivably could be involved in the

* Dr. Summers, a social psychologist, is Associate Professor of Psychology at the University of Kansas, Lawrence, Kansas 66044. Mr. Faucher and Ms. Chapman are doctoral candidates at the University of Kansas. The research reported here was supported by NIMH Grant MH-17621-01 and by NIMH Grant MH-16437-02. The authors gratefully acknowledge the assistance of Karene Will, Bertha Ramsey, and David Purdy.

recognition of treatment needs is obviously broad and would include those who routinely encounter socially deviant behavior as well as those who do not. The present study therefore included samples from several roles, both within (psychiatric aides and clinical trainees) and outside of mental health settings (attorneys and law enforcement officials). In addition, judgments were obtained from disadvantaged members of the black minority, as well as from a haphazard sample of whites, none of whom fell into the role classifications noted above. In all cases the judgment processes of these "nonprofessionals" were compared with those of a sample of professionally trained diagnosticians. In this respect, comparison was made in terms of the following: (1) types of information (cues) utilized when making judgments; (2) number of cues used (scope of cue utilization); (3) judgmental reliability; (4) magnitude of judgments; and (5) judgmental confidence.

METHOD

Methodological and Conceptual Framework

Judgment processes were assessed here using the "lens model" approach formulated by Brunswik (1956), and later extended to problems of clinical inference by Hoffman, (1960) and by Hammond and his associates (Hammond, Hursch, & Todd, 1964). Typically this approach involves the following steps: (1) the *S* is presented a series of multidimensional stimulus objects (e.g., test profiles) and asked to make a judgment about each (e.g., degree of pathology); (2) the *S*'s judgments are then correlated with (or regressed on) the stimulus dimensions (or cues). The magnitude and sign of the resulting coefficients can then be used to infer *which* cues the *S* used when making his judgments, and *how* these cues were used. In short, this approach provides the researcher with a linear model of each *S*'s judgment policy in the domain being studied. (Although questions are frequently raised about the applicability of a linear model to clinical judgment processes, most evidence indicates that such models are adequate. For a full discussion of this issue, see Goldberg, 1968.)

Subjects

A total of 114 *S*s were recruited from eastern Kansas and the metropolitan Kansas City area and were categorized (for purposes of analysis) as follows: professionals (Ph.D. psychologists and psychiatrists), clinical trainees, psychiatric aides, disadvantaged blacks, attorneys, law enforcement officers, and "miscellaneous" laymen. (Subjects were assigned to the "attorney" group by virtue of their legal training; five members of this group were at the time serving as judges.) Differences according to age, sex, or education (independent of occupation) were not found in this study and as a result these distinctions are not discussed here.

Judgment Task

In general the judgment task consisted of two steps: (1)

each *S* was given a series of 60 person profiles where each profile presented seven types of behavioral information (cues) about a hypothetical stimulus person; (2) *Ss* were asked to make a judgment about each person's mental health, and to indicate the confidence with which this judgment was made.

In order to construct a judgment task that would be readily understood by the nonprofessional *Ss* who participated in the study, an effort was made to use profile dimensions, or cues, which denoted observable behavior. Thus a large number of scales that represented observable behavior were selected from several psychiatric rating scales (Lorr, Jenkins, & O'Connor, 1955; Aumack, 1962). Redundancies were eliminated, and the remaining scales were further reduced based on the pretest responses of 15 clinical psychologists. As a result of this procedure, seven cues were selected for construction of the person profiles used here: (1) stated feelings of inadequacy; (2) loses temper when others disagree with him; (3) engages in everyday activities with enthusiasm; (4) drunkenness; (5) job difficulty; (6) blames others when things go wrong; and (7) steals.

Each cue, or profile dimension, was presented in the judgment task as a bar graph that could vary from "never" (1), "rarely" (2), "occasionally" (3), "frequently" (4), to "quite often" (5). Using the seven cues described above, 45 different profiles of hypothetical persons were generated. Each profile was constructed by randomly selecting a value for each of the seven cues, thus yielding a judgment task in which the means and standard deviations for the cues were approximately equal, and in which cue intercorrelations did not differ significantly from zero ($\pm .17$). In order to assess judgmental reliability, 15 of the 45 profiles were repeated, therefore producing a total of 60 profiles. These profiles were printed on standard 8 ½ x 11 inch paper (one profile per page) and bound in judgment booklets.

Response Dimensions

Ss were asked to make two responses for each profile. First, *S* was instructed to judge the mental health of a person having the characteristics shown on that profile. This judgment was indicated by marking a 7-point scale ranging from "mentally healthy" (1) to "mentally ill" (7). (The instructions that preceded the profiles in the judgment booklet defined "mentally ill" in very broad terms; i.e., "needing professional treatment.") Second, *S* was asked to indicate his confidence in his judgment of each profile by marking a 7-point scale ranging from "not at all confident" (1) to "highly confident" (7).

RESULTS

Inasmuch as the judgment model utilized in this analysis is linear, it is important to determine initially whether or not such a model is appropriate in the present task. In this respect the multiple correlation (*R*) between each *S*'s judgment and the seven cues indicates the

extent to which the *S*'s judgments represent a linear combination of the profile cue values. High values of multiple *R* indicate that the *S* utilized the cues in a linear manner, and that a linear correlational model of his judgment process is appropriate. Low values of multiple correlation indicate *either* that the *S* utilized the cues according to a nonlinear (configural or curvilinear) rule or that *S* responded to the profiles in an inconsistent manner.

With the exception of the disadvantaged group ($R = .56$), the median *R*s for the groups were not only quite similar, but also fairly high; that is, the median *R*s for the remaining six groups ranged from .72 to .79. These results indicate that for all but the disadvantaged group a linear model of cue utilization can account for a substantial proportion of the total response variance in each sample.

Cue Utilization

Cue utilization indices were obtained by correlating each *S*'s judgments with the values for each profile cue. In order to obtain stable indices the first 10 trials were treated as "practice" trials; as a result the utilization coefficients were based on the last 50 profiles. The coefficients obtained in this manner were transformed into Fisher *Z* coefficients and then analyzed according to a 7 (Groups) \times 7 (Cues) analysis of variance. (As noted earlier, Cue 3 (enthusiasm) was worded in such a way that *S*s would probably use this cue in an inverse manner; that is, the more enthusiasm, the less pathology. Such was not the case for the other six cues. Therefore in order to avoid a spurious cue effect, the values for Cue 3 were inverted.) This analysis revealed significant main effects for cues, $F(1,107) = 28.68, p < .001$ and for groups, $F(1,107) = 9.88, p < .01$. In addition, a significant interaction was found between cues and groups, $F(6,107) = 2.73, p < .05$.

Individual comparisons revealed that over all groups, "steals," "job difficulty," and "feelings of inadequacy" were utilized more than the remaining four cues. Individual comparisons also revealed that the groups effect is attributable to the rather small cue utilization coefficients derived for disadvantaged *S*s.

The significant cues \times groups interaction, however, indicates that more useful information can be obtained by testing for simple main effects for each cue. These latter analyses revealed significant differences among groups in terms of their utilization of the following profile cues: enthusiasm, $F(6,107) = 5.16, p < .01$; job difficulty, $F(6,107) = 5.27, p < .01$; blames others, $F(6,107) = 3.21, p < .01$; and steals, $F(6,107) = 5.08, p < .01$. In short, it was found that the groups \times cues interaction is attributable to significant differences among the groups on these four cues (see Table 1). Individual comparisons for each finding are reported in turn.

TABLE 1
Mean Utilization Coefficients For Cues On Which
Group Differences Were Obtained

	Cues			
	Enthusiasm	Job Difficulty	Blames Others	Steals
Professionals	.16	.33	.12	.33
Clinical Trainees	.16	.46	.01	.34
Disadvantaged	-.06	.10	.12	.09
Miscellaneous Laymen	.11	.32	.16	.41
Attorneys	-.07	.19	.16	.50
Law Enforcement Officers	-.05	.17	.21	.46
Psychiatric Aides	.03	.22	.23	.41

Enthusiasm Scheffe's (1953) test for individual comparisons revealed that the mean utilization of this cue by the *professional* and *clinical trainee* groups was significantly ($p < .05$) greater than that of the remaining five groups.

Job difficulty Scheffe's test indicated that the mean utilization of "job difficulty" by the *professionals*, *clinical trainees*, and *laymen* was significantly ($p < .05$) greater than that of the remaining four groups.

Blames others Individual comparisons revealed only one statistically reliable ($p < .05$) difference; that is, the *clinical trainees* virtually ignored "blames others," while *Ss* in the other groups tended to place small, though consistent, weight on this cue.

Steals Scheffe's test revealed that the *laymen*, *officers*, *attorneys*, and *psychiatric aides* placed a greater weight on this cue ($p < .05$) than did the *disadvantaged*, *professionals*, and *clinical trainees*.

Scope of Cue Utilization

Scope was measured by the total number of cues *S* utilized significantly ($p < .05$) when making judgments—in this case, the total number of utilization coefficients for each *S* which equaled or exceeded .24. Using this criterion, it was found that very few *Ss* in the entire sample utilized all the information available to them when making their

judgments. Specifically, most Ss (68%) utilized either two or three cues when making judgments. Very few Ss used only one cue (14%), or more than three (18%).

In order to determine whether scope of cue utilization was related to group membership, Ss in each group were categorized as having either a "broad scope" (utilization of three or more cues) or a "narrow scope" (utilization of fewer than three cues). A chi-square analysis of the resulting bivariate frequency distribution failed to yield a significant association between groups and scope of cue utilization ($\chi^2 = 6.56, df = 7, NS$).

Judgment Reliability

Reliability was assessed by correlating each S's responses to the two sets of 15 profiles repeated in the judgment booklet. The mean reliability coefficients for the seven groups are as follows: professionals (.75), attorneys (.69), clinical trainees (.68), law enforcement officers (.62), laymen (.62), psychiatric aides (.51), and disadvantaged (.38). In order to determine whether these group means differed significantly, the reliability indices were transformed to Fisher Z coefficients and subjected to an analysis of variance. This analysis indicated that judgment reliability did indeed vary according to groups, $F (df = 6, 107) = 2.97, p < .05$. Moreover subsequent individual comparisons revealed that this finding can be attributed to the contrast between the disadvantaged and psychiatric aides, on the one hand, and the remaining five groups, on the other. (The latter five groups did not differ significantly.)

Magnitude of Judgments

In order to determine whether or not the groups studied here differed in terms of the *overall* level of pathology they attributed to the persons represented by the profiles, the mean judgment was calculated for each S. It was found that the seven groups were quite similar on this measure: psychiatric aides (4.8), clinical trainees (4.8), professionals (4.7), law enforcement officers (4.6), disadvantaged (4.5), laymen (4.5), and attorneys (4.4). A one-way analysis of variance performed on these mean judgments indicated that these differences can be attributed to chance; that is, $F (df = 6, 107) = .81, NS$.

Confidence

An analysis of the confidence ratings revealed that the Ss were on the whole quite confident of their judgments, despite the somewhat limited amount of information available about each person to be judged; that is, for all groups the mean confidence rating was greater than the midpoint of the 7-point confidence scale: law enforcement officers (5.6), attorneys (5.3), laymen (4.9), professionals (4.8), psychiatric aides (4.8), disadvantaged (4.5), and clinical trainees (4.4). Despite an in-

dication that the police officers and attorneys tended to be more confident in their judgments than the other groups, an analysis of variance failed to yield a significant difference among the groups; $F(df = 6,107) = 2.01, NS$.

DISCUSSION

The nonprofessional Ss studied here differed very little from the professional Ss in terms of either reliability, scope of cue utilization, magnitude of judgments, or judgmental confidence. In this regard the only exception concerned the low reliability indices obtained for the disadvantaged and psychiatric aide groups. It should be noted that these latter groups were characterized by the lowest educational level in the sample. It is possible, therefore, that either the instructions or certain aspects of the judgment task may have been unclear to these Ss, thus contributing to their inconsistent judgments.

On the other hand substantial differences between occupational groups were obtained in terms of their use of profile cues when making judgments; that is, the groups differed in terms of judgment policy. Of particular interest are the differential cue weightings obtained for "enthusiasm" and "stealing." With respect to "enthusiasm," it is important to note that this was the only cue that yielded information about "positive," or adaptive behavior. Although none of the groups relied heavily on this cue (see Table 1), a greater proportion of the professional diagnosticians and clinical trainees (29%) utilized this cue than did the other Ss (12%). This finding suggests that virtually all of the nonprofessional Ss conceived of mental health and mental illness primarily in terms of deviant, or socially disapproved, behavior, with little regard for the area of adaptive functioning that was represented in the judgment profiles.

The significant difference among groups obtained in terms of utilization of "steals" is also noteworthy. Although all groups assigned considerable weight to frequency of stealing when making judgments, a greater proportion of the nonprofessional sample (76%) utilized this cue than did the professionals and clinical trainees (65%). Indeed, 42% of the nonprofessionals (as compared with 29% of the professionals and clinical trainees) assigned a greater weight to this cue than to any of the others. This finding suggests that for the large majority of the nonprofessional Ss, violation of legal norms is an important aspect of their conception of mental illness. This type of behavior appears to be somewhat less important to persons who have had (or are now receiving) formal training in mental health.

Aside from the apparent distinction between professional and nonprofessional conceptions of mental illness suggested by these findings, it should be noted that the differences in judgment policies observed here have implications for potential conflict. As shown elsewhere, differences in judgment policy are an important source of interpersonal conflict; it

might be added that such conflict is typically difficult to resolve (Hammond, 1965; Summers, 1968).

In the present study the differences in judgment policy noted above are reflected in rather low levels of agreement among the Ss. Specifically, when the professional Ss were compared with the other subject groups, the average interjudge correlation was only slightly better than chance: .41 (clinical trainees); .45 (laymen); .41 (attorneys); .41 (police officers); .41 (psychiatric aides); and .20 (disadvantaged). It should be pointed out, of course, that when these correlations are corrected for attenuation, the magnitude of average agreement increases: from .38 (disadvantaged) to .66 (laymen and psychiatric aides). Even when inconsistency or unreliability of judgments is taken into account, however, it is clear that the agreement between the professional and nonprofessional Ss is less than complete.

It should be emphasized, of course, that the sample studied here was quite limited, both in terms of size and regional representation. As a result these findings, though statistically significant, should be regarded as only tentative. Despite these limitations, this study provides additional support for the utility of a multiple-cue correlational model as a means of studying mental health judgments. As illustrated here this approach yields quantitative information about judgment processes, and can allow predictions about the magnitude and locus of disagreement between individuals who are making judgments about the mental health of others. Under conditions in which individuals from diverse ethnic, educational, and occupational backgrounds play a role in mental health settings, such information would appear to be essential.

REFERENCES

- Aumack, L. A social adjustment behavior rating scale. *Journal of Clinical Psychology*, 1962, 18, 436-441.
- Brunswick, E. *Perception and the representative design of psychological experiments*. Berkeley: University of California Press, 1956.
- Ellsworth, R. *Nonprofessionals in psychiatric rehabilitation*. New York: Appleton-Century-Crofts, 1968.
- Ewalt, P. *Mental health volunteers: The expanding role of the volunteer in hospital and community mental health services*. Springfield, Illinois: Charles C. Thomas, 1967.
- Goldberg, L. Simple models or simple processes? Some research on clinical judgments. *American Psychologist*, 1968, 23, 483-496.
- Hammond, K. R. New directions in research on conflict resolution. *Journal of Social Issues*, 1965, 21, 44-66.
- Hammond, K. R., Hursch, Carolyn, & Todd, F. J. Analyzing the components of clinical inference. *Psychological Review*, 1964, 71, 438-456.
- Hoffman, P. The paramorphic representation of clinical judgment. *Psychological Bulletin*, 1960, 57, 16-131.
- Hunt, W. A., Jones, N. F., & Hunt, Edna. Reliability of clinical judgments as a function of clinical experience. *Journal of Clinical Psychology*, 1957, 13, 377-378.
- Lorr, M., Jenkins, R., & O'Connor, J. Factors descriptive of psychopathology and behavior of hospitalized psychotics. *Journal of Abnormal and Social Psychology*, 1955, 50, 78-86.
- Marabian, A. & Reed, H. Factors influencing judgments of psychopathology. *Psychological Reports*, 1969, 24, 323-330.

- Matarazzo, J. D. Some national developments in the utilization of nontraditional mental health manpower. *American Psychologist*, 1971, 26, 363-372.
- Scheffe', H. A. A method for judging all possible contrasts in the analysis of variance. *Biometrika*, 1953, 40, 87-104.
- Summers, D. A. Conflict, compromise, and belief change in a decision-making task. *Journal of Conflict Resolution*, 1968, 12, 215-221.
- Yarrow, M., Schwartz, C., Murphy, H., & Deasy, L. The psychological meaning of mental illness in the family. *Journal of Social Issues*, 1955, 11, 12-24.