Variables Affecting Psychiatrists' and Teachers' Assessments of the Dangerousness of Mentally Ill Offenders

Vernon L. Quinsey and Rudolf Ambtman
Mental Health Centre, Penetanguishene, Ontario, Canada

Nine high school teachers and four forensic psychiatrists rated the psychiatric assessments, histories, and offense descriptions of 9 child molesters, 10 property offenders, and 11 serious offenders against adults. The three types of patient data were rated separately and together on the likelihood of a property offense, the likelihood of an assaultive offense, and the seriousness of an assault should one occur, if these male patients were to be released from a maximum security mental hospital. Raters also judged whether the patients should be released. There was very low interrater congruence in all ratings and in the release decisions for both occupational groups, although the correlations between the occupational group averages were very high. Ratings of the combined information were well predicted from a linear combination of the ratings of its three components for both rater groups. Psychiatric assessment data made little contribution to the appraisal of the overall file. These data question the expertise of forensic psychiatrists as well as the usefulness of psychiatric assessment data in the prediction of dangerousness.

Assessment of the dangerousness of mentally ill offenders are made by psychiatrists in order to advise the courts regarding these offenders' dispositions and to determine when a mentally ill offender should be released from a psychiatric institution. In both instances, the psychiatrist's expert opinion has a profound effect on the life of the person being assessed, particularly inasmuch as mentally ill offenders are often detained in maximum security facilities on fully indeterminate bases.

The research literature on the prediction of dangerousness suggests that these predictions are conservative and of low accuracy (Megargee, 1976; Monahan, 1975; Shah, 1978; Wenk, Robison, & Smith, 1972). This interpretation, however, rests largely on the results of follow-up studies of released mentally disordered offenders. Unfortunately, interpretation of these studies is ambiguous because very few such patients are released when psychiatrists judge them to be serious dangers to others and the dangerousness of these kept incarcerated is not known (Quinsey, Pruesse, & Ambtman, Note 1). The Baxstrom study (Steadman & Cocozza, 1974) of mental patients released from maximum security because of a judicial decision is an exception in that no psychiatric discretion was exercised. This study found that very few of these patients were dangerous. Unfortunately, however, the average age of the follow-up sample was quite advanced at the time of release. Follow-up data from other maximum security mental hospitals (e.g., Pruesse & Quinsey, 1977) indicate that a considerable proportion of those released are dangerous to others even after a psychi-
atric screening. Follow-up studies which have attempted to evaluate clinical determinations of dangerousness (e.g., Kozol, Boucher, & Garofalo, 1972) have suffered from the methodological problem of not specifying interjudge reliabilities for the assessments in addition to the problem of differential release policies alluded to above.

In view of the important influence that the forensic psychiatrist's expertise has on the destiny of offenders suspected of being mentally ill and dangerous, it behooves us to carefully examine the nature of this expertise as it has been challenged, albeit indirectly, by the research literature (e.g., Quinsey, 1977; Quinsey et al., Note 1). A study in point (Quinsey & Ambtman, Note 2) examined ratings of men found not guilty by reason of insanity or unfit for trial made by three psychiatrists and a psychologist at 105 conferences in which the patient's possible release or transfer was considered. A modest amount of interrater agreement on the dangerousness of the patients was obtained (average $r = .60$). An average multiple $R$ of .48 was found between the dangerousness ratings and 10 patient characteristics obtained from the clinical files. When the raters were asked to rank the 10 variables in order of importance, they agreed with neither each other nor the results of their own regression analyses. Unfortunately, however, there are certain ambiguities which arise in the interpretation of these data because (a) it was impossible to determine on what information the ratings were actually based despite the effort made to obtain comprehensive data from the files and (b) the fact that the raters discussed each case before making their ratings presumably maximized rater congruence.

In view of the limitations in the above study and in this research literature generally, a convincing examination of psychiatric expertise in the prediction of dangerousness remains to be accomplished despite the long history of clinical judgment studies (Taft, 1955; Wiggins, 1973). The present study attempts such an examination by determining whether psychiatrists' assessments of the dangerousness of mentally ill offenders met four criteria for expert judgment. In order for a group of professionals to demonstrate expertise in a given field they must (a) agree amongst themselves in their judgments, (b) be accurate in their judgments (this accuracy will be limited by the amount of their disagreement), (c) make different judgments from laypersons, and, as a corollary of c, (d) make use of the specialized procedures they employ in arriving at a judgment.

The strategy of the present study was to compare psychiatrists' and teachers' ratings of the dangerousness of three types of patients defined according to offense type: child molesters, serious offenders against persons (mostly murderers), and property offenders. Teachers were chosen as laypersons because they were literate but not trained in psychiatry or law. The last two patient groups were chosen to maximize differences in the severity of their admission offense and hence probable dangerousness ratings; the child molesters were included because they pose special problems in assessment associated with the lack of opportunity to commit a new offense while incarcerated and public abhorrence of the act itself (Quinsey, 1977).

In contrast to our earlier study (Quinsey & Ambtman, Note 2), the ratings were made without the benefit of previous discussion, and the amount and type of information available to the raters was controlled. This procedure allowed us to examine to what extent each type of information (history, offense, and assessment) was used in forming an overall appraisal.

Method

Raters

Five female and four male public secondary school teachers were each paid $15.00 an hour to perform the ratings. The teachers ranged in age from 25 to 56 years and none had formal training in psychiatry, clinical psychology, or law beyond an introductory course in psychology.

A group of four male psychiatrists served as volunteer raters. Three of the psychiatrists worked at the maximum security Oak Ridge Psychiatric Institution and one at an Ontario correctional center. All were very experienced in the treatment
and assessment of mentally disordered offenders and considered to be experts in the area of forensic psychiatry. Each had had extensive court experience as well.

**Rated Patients**

A total of 30 patients were selected for the study according to whether they had committed an offense leading to admission which fell into only one of the following categories: (a) a nonsexual assaultive offense against an adult (in most cases a murder), (b) a sexual offense involving bodily contact with a child 13 years of age or younger, or (c) an offense against property. Eleven patients were offenders against adults, 9 were child molesters, and 10 had committed a property offense. These patients were the first found in a backward search of the files which met the criteria for inclusion in the respective groups. The search began with patients who were in Oak Ridge in 1973, 3 years before this study began. The unequal ns of the groups was the result of a coding error.

**Rater Materials**

A list of the variables to be included from the files was drawn up in consultation with two of the four psychiatrists who were used as raters. The list included all the variables contained in the files which the psychiatrists indicated would be important in arriving at an assessment of patient dangerousness. All the relevant data mentioned in the file were included in specially prepared summaries which were used for the purpose of rating.

Three separate summaries were prepared for each file. Each summary contained the patient's age and the number of years he had been in Oak Ridge as of August 1, 1973. This information was placed in each summary because these data are always available whenever an assessment of patient dangerousness is made, regardless of what type of data was thought to be most important by the psychiatrist. The first type of summary contained only information which was descriptive of the offense(s) leading to admission; the second contained a description of the patient's history prior to the offense leading to admission, including his family background, previous offenses, history of institutionalization, and the like; the third summary contained assessment data obtained since the time of the patient's relevant admission—for example, psychological test data (MMPI, IQ, and Rorschach summaries where available), mental statuses (the patient's interview demeanor, orientation for time, place, and person, diagnostic impression, etc.), progress notes (ward behavior, work records, security status, medications, etc.), but no information pertaining to the patient's offense or history.

Each patient was rated by each rater on two occasions. On the first occasion, the three summaries were rated separately for half of the patients, whereas the combined information (the "total file") was rated for the remainder; on the second occasion, the remaining material for each patient was rated. The order in which the information was presented over rating occasions was randomly alternated over patients and was different for each rater. The order in which patients were rated on each rating occasion was randomly determined and different for the various raters. The rating occasions were separated by a minimum of 5 weeks, so the raters would not remember which summaries went together; in addition, it was implied that the raters were to rate 60 different patients.

Because some of the three Oak Ridge psychiatrists would have had contact with some of the patients, the following steps were taken to prevent the real patients' identities from being recognized: All the dates were moved back 10 years, no names were used, geographical locations were changed or kept vague, and unusual details which would be likely to be remembered were made vague or slightly altered. In this connection it is relevant that the most recent patient data included were over 3 years old at the time of the study.

Subjects were asked to rate on 9-point scales how likely they thought the patient was, if released to the street at the time of assessment, to commit (a) a property offense within 12 months and (b) an assaultive offense against persons within 12 months. In addition, raters rated their confidence in each rating. Subjects also rated on a 9-point scale ranging from "not causing injury" to "causing death" how serious the offense against persons would be should the patient commit one. Last, the subjects indicated whether they thought the man should be released at the time of the assessment.

**Results**

**Interrater Reliabilities**

The interrater reliability data are presented in Table 1. As shown in the table, the interrater correlations within the occupational groups were quite low, the highest being .675 for the teachers’ prediction of the seriousness of an assaultive offense should one occur based upon the patients’ offense data. In general, the interrater reliabilities were higher for the offense category and history and lowest for the assessment category for both groups of raters. The psychiatrists did not show higher rates of agreement than the teachers. The average teacher intercorrelations were higher in 10 out of the 12 comparisons.

Table 1, bottom section, also shows that the teachers and psychiatrists tended to rate
VERNON L. QUINSEY AND RUDOLF AMBTMAN

Table 1

Average Interrater Reliabilities

<table>
<thead>
<tr>
<th>Item</th>
<th>Total</th>
<th>Offense</th>
<th>History</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychiatriests (n = 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood of property offense</td>
<td>.484 (.226)</td>
<td>.616 (.053)</td>
<td>.572 (.167)</td>
<td>.309 (.178)</td>
</tr>
<tr>
<td>Likelihood of assaultive offense</td>
<td>.185 (.182)</td>
<td>.344 (.109)</td>
<td>.505 (.179)</td>
<td>.145 (.202)</td>
</tr>
<tr>
<td>Seriousness of assaultive offense</td>
<td>.382 (.156)</td>
<td>.498 (.106)</td>
<td>.331 (.136)</td>
<td>-.097 (.154)</td>
</tr>
<tr>
<td>Teachers (n = 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood of property offense</td>
<td>.459 (.109)</td>
<td>.619 (.120)</td>
<td>.574 (.105)</td>
<td>.132 (.237)</td>
</tr>
<tr>
<td>Likelihood of assaultive offense</td>
<td>.240 (.188)</td>
<td>.463 (.157)</td>
<td>.507 (.128)</td>
<td>.276 (.151)</td>
</tr>
<tr>
<td>Seriousness of assaultive offense</td>
<td>.574 (.154)</td>
<td>.675 (.129)</td>
<td>.336 (.176)</td>
<td>.207 (.180)</td>
</tr>
</tbody>
</table>

Average ratings for teachers and psychiatrists (N = 13)

| Likelihood of property offense | .876 | .889 | .803 | .194 |
| Likelihood of assaultive offense | .550 | .726 | .833 | .705 |
| Seriousness of assaultive offense | .816 | .786 | .658 | .618 |

Note. The numbers in parentheses are standard deviations. Correlations of .361 and .463 are significant at the .05 and .01 levels of significance, respectively.

the patients similarly. The correlations between the average ratings of the teachers and the average ratings of the psychiatrists were all significant except for the prediction of the likelihood of a property offense based on assessment data.

It may be argued that the dimensions discussed above (i.e., likelihood of a property offense, likelihood of an assaultive offense, and seriousness of an assaultive offense should one occur) are not the only dimensions used by the psychiatrists in deciding whether a person is fit for release and, therefore, that the psychiatrists may agree quite well on the actual release decision although they show low reliabilities on the dimensions already discussed. The interrater agreements on whether to release a patient are, therefore, very important. These data are shown in Table 2. First the percent agreements were calculated for each pair of raters and then averaged; second, the kappa statistic, (Cohen, 1960), the proportion of cases in which the judges agreed after correcting for chance, was similarly averaged over pairs of raters. In cases in which only one cell in the 2 x 2 agreement matrix was filled, kappa was assigned a value of unity. As shown in

Table 2

Average Agreement Among Pairs of Raters on the Release Decisions

<table>
<thead>
<tr>
<th>Information type</th>
<th>Rater</th>
<th>No. pairs</th>
<th>Total</th>
<th>Offense</th>
<th>History</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average % agreement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatrists</td>
<td>6</td>
<td>75.56 (.737)</td>
<td>66.67 (9.03)</td>
<td>53.33 (15.15)</td>
<td>63.33 (11.06)</td>
<td></td>
</tr>
<tr>
<td>Teachers</td>
<td>36</td>
<td>82.78 (12.51)</td>
<td>67.59 (13.11)</td>
<td>67.78 (12.12)</td>
<td>61.48 (13.55)</td>
<td></td>
</tr>
<tr>
<td>Teachers and psychiatrists</td>
<td>36</td>
<td>80.00 (11.36)</td>
<td>69.63 (11.88)</td>
<td>59.91 (11.10)</td>
<td>65.19 (14.04)</td>
<td></td>
</tr>
<tr>
<td>Average kappa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatrists</td>
<td>6</td>
<td>.055 (.059)</td>
<td>.238 (.162)</td>
<td>.154 (.156)</td>
<td>.097 (.148)</td>
<td></td>
</tr>
<tr>
<td>Teachers</td>
<td>36</td>
<td>.106 (.296)</td>
<td>.154 (.208)</td>
<td>.323 (.173)</td>
<td>.215 (.180)</td>
<td></td>
</tr>
<tr>
<td>Teachers and psychiatrists</td>
<td>36</td>
<td>.147 (.301)</td>
<td>.248 (.226)</td>
<td>.196 (.136)</td>
<td>.230 (.185)</td>
<td></td>
</tr>
</tbody>
</table>

Note. The numbers in parentheses are standard deviations.
Table 2, although the percent agreements were higher for both psychiatrists and teachers when given the total file on which to make a judgment, rather than on its separate components, the kappa coefficient was very small, indicating that the percent agreements were inflated by chance. Neither psychiatrists nor teachers, therefore, exhibited very high levels of agreement on which patients should be released.

**Mean Differences in Ratings**

Although the correlations between the psychiatrists and teachers shown in Table 1 are quite high, it is possible that the two rater types differed in their mean ratings; for example, that the rater groups may have differed in their degree of conservatism. To assess this possibility, analyses of variance were computed for each of the rated dependent variables: likelihood of a property offense, likelihood of an assaultive offense, and seriousness of an assaultive offense should one occur. For each analysis, in order to achieve equal n, patients were randomly discarded so that each patient group contained nine members and the averages of the occupational groups' ratings were employed in the analyses. The variables were: rater type (psychiatrist or teacher), patient type (offender against adult, child molester, or property offender), patients within rater type, and information type (total, assessment, history, and offense). Scheffé tests were computed for significant main effects and interactions with the alpha level set at 10% (Myers, 1966). Only the F values for contrasts reaching Scheffé's adjusted criterion are reported. A summary of the analyses of variance is shown in Table 3.

Considering the first dependent measure, the psychiatrists rated the patients as more likely to commit a property offense than did the teachers. Rater type also interacted significantly with information type but subsequent Scheffé contrasts were not significant. There was a main effect of patient type because, not surprisingly, property offenders were viewed as more likely to commit a future property offense. The significant information type effect reflected the finding that ratings of likelihood of a property offense made from the total file and patient history were higher than ratings of the remaining information types, $F(1, 72) = 42.42$, $p < .001$.

Turning to the ratings of the likelihood of a future assault, the effect of patient type occurred because the child molesters were rated as more likely to commit a future as-

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Likelihood property</th>
<th>Likelihood assault</th>
<th>Seriousness assault</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater type (A)</td>
<td>1</td>
<td>34.91****</td>
<td>0.56</td>
<td>4.00</td>
</tr>
<tr>
<td>A × Patient Type (B)</td>
<td>2</td>
<td>1.35</td>
<td>13.76****</td>
<td>2.70</td>
</tr>
<tr>
<td>A × Patients (C) within B</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>10.05****</td>
<td>7.09***</td>
<td>9.85****</td>
</tr>
<tr>
<td>C within B</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information type (D)</td>
<td>3</td>
<td>15.56****</td>
<td>8.36****</td>
<td>14.92****</td>
</tr>
<tr>
<td>B × D</td>
<td>6</td>
<td>4.82****</td>
<td>4.73****</td>
<td>6.66****</td>
</tr>
<tr>
<td>C × D within B</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A × D</td>
<td>3</td>
<td>3.11*</td>
<td>9.80****</td>
<td>3.71**</td>
</tr>
<tr>
<td>A × B × D</td>
<td>6</td>
<td>1.33</td>
<td>1.22</td>
<td>1.34</td>
</tr>
<tr>
<td>A × C × D within B</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p = .05$.
** $p = .025$.
*** $p = .005$.
**** $p = .001$. 

Table 3

* Analysis of Variance for the Three Rated Variables
saultive offense than the other patient types, $F(1, 24) = 12.17, p < .005$. The main effect of information type reflected the higher likelihood of assault ratings given to the total file than the other categories of information, $F(1, 72) = 24.32, p < .001$.

The significant Rater Type × Patient Type interaction occurred because teachers rated offenders against adults and child molesters as more likely to commit an assault than the psychiatrists did but rated property offenders as less likely, $F(1, 24) = 12.79$.

Figure 1. Average percentage of patients released and mean seriousness of assault ratings for teachers and psychiatrists by type of information rated.
The interaction between rater type and information type resulted from the psychiatrists rating patients as more likely to assault when given only assessment data but the teachers giving higher ratings to the other three information types, $F(1, 72) = 58.48, p < .001$.

Perhaps the most interesting effect in the likelihood of assault ratings is indicated by the Patient Type $\times$ Information Type interaction. In part this interaction occurred because the history of the assaultive offenders against adults received very low likelihood of an assault ratings, whereas their offense data received very high ratings; this difference in type of information was much smaller for the other two patient groups, $F(1, 72) = 35.33, p < .001$.

The seriousness ratings are shown on the right side of Figure 1. The patient type effect reflected the expected higher ratings of the offenders against adults in comparison to the other patient groups, $F(1, 24) = 20.53, p < .001$. The information type effect and Information $\times$ Patient Type interaction both resulted from the very high seriousness ratings given the offenders against adults when their offense descriptions or total files were considered, $F(1, 72) = 144.74, p < .001$. The significant Rater Type $\times$ Information Type interaction resulted from the finding that teachers gave higher seriousness ratings to all information types than psychiatrists, save on the assessment data, $F(1, 72) = 23.31, p < .001$.

### Mean Percent Released

As shown on the left side of Figure 1, the average percentage of patients who were recommended for release by the two occupational groups corresponded rather well. It is noteworthy that these percentages correspond very closely to the mean seriousness of assault ratings for the offender against person patient type. When the raters decided whether an offender against persons should be released or rated the seriousness of his potential assaultive offense, they viewed the matter very differently according to whether they based their judgments on the total file or offense data or whether they employed assessment or history data. The other feature of the data depicted by the graph is that fewest patients were released and the ratings of seriousness were highest when the total file was examined. The higher interrater percent agreements and the low kappas for the release decision based on the total file as opposed to the other information types appears to be due to the relatively conservative release decisions based on the total file, which inflated the probability of chance agreement.

### Relative Contribution of Types of Information

Stepwise regression analyses were computed for the averages of the teachers’ and psychiatrists’ ratings and for each psychiatrist separately for the three rated variables. In each analysis, the rating of the total file was predicted from the patient ratings based on offense, history, and assessment. These data are shown in Table 4. With the exception of the likelihood of assault ratings made by the psychiatrists, the multiple $R^2$s for the rater group averages were very high, indicating that the appraisal of the total file was well predicted from a linear combination of its components. The usefulness of each information type was computed following Darlington (1968), where the usefulness of a variable is defined by the amount that $R^2$ would drop if the variable in question were deleted from the analysis and the others appropriately reweighted. It can be seen in the table that assessment data never contributed very much to the overall $R$ for either group of raters and that for the prediction of the seriousness of an assaultive offense, previous offense description was the only important variable.

### Discussion

#### Psychiatric Expertise in the Prediction of Dangerousness

The purpose of this study was to examine the nature of psychiatric expertise in the assessment of the dangerousness of mentally ill offenders.
Table 4  
Predicting the Ratings of the Total File From Ratings of Its Components

<table>
<thead>
<tr>
<th>Rater</th>
<th>Beta</th>
<th>Offense</th>
<th>History</th>
<th>Assessment</th>
<th>R</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Likelihood of property offense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers' average</td>
<td>.486 (.19, .001)</td>
<td>.596 (.28, .001)</td>
<td>.006 (0, ns)</td>
<td>.921</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Psychiatrists' average</td>
<td>.667 (.17, .001)</td>
<td>.540 (.18, .001)</td>
<td>-.353 (.05, .05)</td>
<td>.845</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Psychiatrist</td>
<td>1</td>
<td>.132</td>
<td>.173</td>
<td>-.038</td>
<td>.251</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.307</td>
<td>.448</td>
<td>.278</td>
<td>.840</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.402</td>
<td>.392</td>
<td>.166</td>
<td>.740</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.517</td>
<td>.239</td>
<td>-.425</td>
<td>.634</td>
<td>.005</td>
</tr>
<tr>
<td><strong>Likelihood of assaultive offense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers' average</td>
<td>.472 (.22, .001)</td>
<td>.603 (.36, .001)</td>
<td>.382 (.15, .001)</td>
<td>.853</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Psychiatrists' average</td>
<td>.190 (.03, ns)</td>
<td>.393 (.15, .05)</td>
<td>.161 (.02, ns)</td>
<td>.471</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Psychiatrist</td>
<td>1</td>
<td>-.085</td>
<td>.150</td>
<td>.044</td>
<td>.183</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.288</td>
<td>.612</td>
<td>.095</td>
<td>.631</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.005</td>
<td>.555</td>
<td>-.200</td>
<td>.564</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.211</td>
<td>-.044</td>
<td>.075</td>
<td>.203</td>
<td>ns</td>
</tr>
<tr>
<td><strong>Seriousness of assaultive offense</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teachers' average</td>
<td>.880 (.77, .001)</td>
<td>.090 (.01, ns)</td>
<td>.171 (.03, .05)</td>
<td>.906</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Psychiatrists' average</td>
<td>.741 (.48, .001)</td>
<td>.031 (0, ns)</td>
<td>.299 (.09, .025)</td>
<td>.784</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Psychiatrist</td>
<td>1</td>
<td>.558</td>
<td>.017</td>
<td>.103</td>
<td>.591</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.682</td>
<td>.076</td>
<td>-.082</td>
<td>.665</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>.414</td>
<td>-.153</td>
<td>.169</td>
<td>.463</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>.019</td>
<td>-.069</td>
<td>.125</td>
<td>.136</td>
<td>ns</td>
</tr>
</tbody>
</table>

*Note.* The numbers in parentheses are: first, the amount that $R^2$ would drop if the variable were removed, and second, the significance of the component.

Disordered offenders. The first criterion for expertise was that the experts agree amongst themselves in their ratings; their highest average interrater correlation for the total file was .482, where they predicted the likelihood of a property offense. The second criterion was that they be accurate in their judgments. It is clear from the low interrater reliabilities that the validity coefficients, if computed (i.e., if we had follow-up data), would be very low, as they cannot be higher than the interrater reliabilities.

The third criterion was that the psychiatrists make different judgments from laypersons. The correlations among the average teacher and psychiatrist ratings indicate, with the exception of their ratings of the assessment data when predicting the likelihood of a property offense, that their ratings were very similar. Furthermore, the regression analyses indicated that they combine and weight the three information types similarly in arriving at a rating of the overall file. As would be expected from these similarities in ratings, the average percentage of patients for whom release was recommended by the two occupational groups varied closely together over patient type and information type.

There is, nevertheless, some support for the idea that psychiatrists differ from teachers in their mean ratings, to be found in the analyses of variance. Psychiatrists saw patients as more likely to commit a property offense than did the teachers. The psychiatrists rated offenders against adults and child molesters as less likely to commit an assault and property offenders as more likely
than did the teachers. The psychiatrists also
saw patients as more likely to commit an
assault than the teachers saw them when only
assessment data were considered but less
likely otherwise. A similar effect occurred
in the seriousness ratings. Psychiatrists,
therefore, did show certain differences from
teachers in their mean ratings. The question
of whether the psychiatrists were more ac-
curate in their ratings than the teachers,
however, is vitiated by the low interrater
reliabilities for both occupational groups.

The fourth criterion was that psychiatrists
employ specialized assessment techniques in
arriving at a judgment, as it would be ex-
pected that assessment data would be an im-
portant determinant of their appraisal of the
total file. The regression analyses, however,
indicate that neither the psychiatrists nor
teachers gave much weight to the assessment
data in arriving at a judgment. The only ex-
ception to this finding was that the psychi-
atrists used the assessment data in predicting
the likelihood of a property offense. How-
ever, the beta weight for this variable was
negative, making the interpretation of the
meaning of its contribution difficult (Dar-
lington, 1968).

In summary, these data indicate that none
of the four criteria for expert judgment was
met. However, a plausible criticism of this
conclusion is that if more information had
been provided for the psychiatrists (partic-
ularly an interview with the patient), their
interrater agreement would improve. This
criticism is implausible for two reasons:
First, as shown in Table 1, the interrater rs
did not increase with more information; and
second, in our earlier study (Quinsey &
Ambtman, Note 2), the interrater agreement
on dangerousness was modest even when the
psychiatrists interviewed the patient and dis-
cussed the case before or during the time
they made their ratings.

Influence of the Type of Information
Considered

When the psychiatrists and teachers con-
sidered the total file, they tended to view the
patients as more likely to commit property
and assaultive offenses, to judge that if an
assaultive offense were to occur that it
would be more serious, and to be less likely
to recommend that the patient be released
than when they considered the other types of
information. One interpretation of this find-
ing is that the raters were looking for signs
that the patients were dangerous and were
more likely to find such an indicator when
they were given more information—that is,
the more they knew, the more conservative
they tended to be.

It was of interest that the type of informa-
tion considered had the greatest influence on
the seriousness ratings of men who had com-
mitted serious offenses against adults. When
the total file or offense descriptions were
used as the basis of the ratings, these patients
were seen as capable of very serious crimes;
whereas, when only the histories or psychi-
atrific assessments were considered, these pa-
tients were not seen as very capable of com-
mitting serious crimes against persons. In
fact, as shown in Figure 1, when only his-
tories were considered, the offenders against
adults were the most likely to receive a re-
lease recommendation. This finding fits in
very well with Megargee's (Megargee, Cook,
& Mendelsohn, 1967) view that persons who
commit very serious crimes against persons
tend to be those who have led normal, even
timid, lives in the community and to be
unlikely to have had previous criminal in-
volvement. The raters in the present study
presumably found very little antisocial or
criminal material in the offenders against
adults' histories and therefore recommended
release.

When the appraisals of the complete file
were predicted from the various information
types, it was noteworthy that the psychiatric
assessment data were never weighted very
heavily by either psychiatrists or teachers for
any of the three variables rated. These data,
therefore, question the usefulness of psychi-
atrific examinations in the prediction of dan-
gerousness, since the psychiatrists themselves
based their appraisal on the patients' his-
tories and offense descriptions. Perhaps psy-
chiatric examinations would be best re-
stricted to determining whether an offender
is treatable and should not address the issue of dangerousness.

Reference Notes


References


Received September 12, 1978