The Reliability and Validity of Dangerous Behavior Predictions

C. D. WEBSTER, Ph D
D. S. SEPEJAK, MA
R. J. MENZIES, MA
D. J. SLOMEN, MA
F. A. S. JENSEN, MBBS
B. T. BUTLER, MD

Taken generally, the literature over the past ten or fifteen years casts considerable doubt on the ability of forensic psychiatrists to predict dangerous behavior of their patients.[1] Much of this work has been based on large-scale studies of persons found not guilty by reason of insanity or incompetent to stand trial.[2,3] Such studies, though extremely important, are limited because they deal only with patients who are alleged to have committed extremely serious crimes. Other problems are that in these previous ventures clinicians have not recorded their predictions in a reasonably standard fashion and that interclinician reliabilities of predictions have been lacking. In an important recent study Mullen and Reinehr[4] put the matter well when, after having found essentially no relationship between clinical predictions of dangerousness and outcome at four years, they stated: "In point of fact, dangerousness has never been demonstrated to be an identifiable personality dimension" and go on to provide the challenge that gives impetus to the present article: "No investigator has been able to show agreement between judges or other instruments which purport to predict it" (p. 230).

In a previous study published in this journal,[5] we offered data based on a sample of 598 persons assessed within the Brief Assessment Unit (BAU) at METFORS. The group-based, court-ordered assessment procedures used in the BAU are described in a recent book.[6] Four psychiatrists made predictions about future dangerousness on a four-point scale, and all patients were followed via searches of records after an interval of two years. The methodology employed in the previous study was similar to that described below for the present project. We discovered in the previous study that aside from previous offense patterns, most demographic variables (age, sex, previous psychiatric history, and so forth) could *not* be used to predict outcomes but that clinical opinion was reasonably effective (though with many false positives).

Webster is Research Scientist at METFORS, Clarke Institute of Psychiatry and Professor of Psychiatry, Psychology and Criminology at the University of Toronto. Sepejak and Menzies were previously Research Scientists at METFORS. The former is now Senior Research Officer at the Ontario Ministry of Correctional Services and the latter is Instructor in the Department of Criminology, Simon Fraser University. Slomen formerly coordinated the prediction of dangerousness project at METFORS and is now a member of the Psychology Department at Queen Street Mental Health Centre. Jensen is Deputy-Director - Clinical at METFORS and Associate Professor of Psychiatry at the University of Toronto. Butler, formerly Staff Psychiatrist at METFORS is now in private practice and with the Forensic Service of the Clarke Institute of Psychiatry. This paper was presented at the Annual Meeting of the American Academy of Psychiatry and the Law in October 1982. Requests for reprints should be addressed to C.D. Webster, METFORS, 1001 Queen Street West, Toronto, Ontario, M6J 1H4.

The previous study, a preliminary description of which is also given in our recent monograph,[6] suffered from two major methodological limitations. The four-point predictor scale was very crude, and no data on the reliability of predictions were offered. In addition, the previous study may have offered an unfair test in that the assessing psychiatrists, at the time of making their evaluations, had no idea their predictions would be checked against actual outcome. In the present study we attempted to replicate our previous findings and to overcome partially the limitations inherent in that work. The present study, though based on a smaller number of subjects, attempted to improve on its predecessor in three main ways: (1) we employed a 7-point rather than a 4-point scale for prediction of "dangerous to others in the future"; (2) we used a pair of specially trained coders to make assessments of dangerousness; and (3) we fully informed the clinical staff their predictions would be validated through follow-up searches of records. The reason for using paid coders in addition to psychiatrists was simple — administratively it would have been too difficult and too expensive to employ forensic psychiatrists to do the evaluative task on an uninvolved basis. Yet in this study we do report data from psychiatrists who undertook the assessments as part of their ordinary clinical duties within the BAU interdisciplinary assessment team.

Method

Subjects 190 men and 27 women were assessed consecutively within the BAU over a four and one-half month period. All persons were remanded by the court. Almost all were at the pretrial stage, and they faced a great variety of charges ranging from failure to appear before the court to murder.

Procedure The usual BAU assessment procedure centers on a group interview. At the time this study was undertaken the interview was normally conducted by the psychiatrists with psychologist, nurse, correctional officer, and social worker in attendance. The actual membership of the team changed from day to day. But for the duration of this study, a research assistant sat in on all group interviews (without participating). Two other research assistants watched the interviews from behind a one-way mirror. These were our external coders. At the conclusion of each interview, and before discussion took place among the clinicians, the research assistant collected from the clinicians and external coders their opinions about dangerousess as expressed on a prepared form. This form was the recording device for a 23-item "Dangerous Behaviour Rating Scheme" (DBRS) devised by the authors in consultation with clinical colleagues[7] and with some influence from Megargee's writings.[8] For this report interest is restricted to a key item, "Dangerousness to Others in the Future." It should however be noted that at the time this study was in progress all BAU clinicians were using the DBRS as a matter of routine, and that prior to the commencement of the project, the two external coders were trained to use the scheme.

Two years after the assessment interview, the records of five major psychiatric hospitals in the proximity were examined to determine whether the previously assessed individuals had been admitted. As well, we searched the correctional service records to find out if there had been new arrests (or misconducts on the

part of persons who were or had been in prison). Details about incidents were extracted from the records and typed on sheets. A given patient (identified only by code number) might have but a single entry or several. Each case was then rated for degree of dangerousness by nine criminology graduate students on an 11-point scale, using a procedure described in our previous report.[5] In the analyses outlined below we use the average obtained from the students.

Results

Inter-Rater Reliability We are here concerned with the reliability of the two external coders on the item "Dangerousness to Others in the Future." The intraclass correlation was +0.72 between the two coders on the 7-point scale. As a check to see whether the scoring was stable over patients we broke the sample into two periods (first 105 versus last 105 subjects) and completed an analysis of variance. Although this yielded a significant between-coder effect the actual mean difference between the two coders was not great (4.78 versus 4.32). It simply indicates one of the coders perceived slightly more potential for dangerous behavior than the other. There was no reliable period effect and no period X rater interaction. The general conclusion is that the two raters were measuring similar qualities while making their assessments of dangerousness to others in the future.

Predictive Ability of Coders and Psychiatrists We found at least one piece of follow-up information on 158 of the 217 former patients (73 percent). It is these patients who form the present sample.[9] Our main interest was to determine the correspondence between prediction and outcome for these two coders and the four psychiatrists. Correlations are summarized in Table 1 where the top two rows show these two raters achieved a moderate degree of predictive success. The performance of the psychiatrists matched that of the coders on average, but it was extremely variable. Psychiatrist 4's predictive performance of +0.49 is the highest we have found in our present series of research studies.

An even simpler way of summarizing the present data is to collapse the predictor and outcome scales and do chi-square tests. The results of such tests are shown in Tables 2, 3, and 4 (next page). We arbitrarily broke the 7-point predictor scale into three intervals by combining scores of 1, 2, and 3, allowing scores of 4 to stand as the mid-point, and combining scores of 5, 6, and 7. Much the same was done with the 11-point outcome scale. The extreme high scores (9, 10,

Table 1. Pearson Prediction-Outcome Correlations
DBRS Item 22 "Dangerousness to Others" — Future

External Coder 1	r = +0.23	n = 140	p = 0.01
External Coder 2	r = +0.18	n = 152	p = 0.01
Psychiatrist 1	r = +0.33	n = 40	p = 0.02
Psychiatrist 2	r = −0.01	n = 28	p = NS
Psychiatrist 3	r = +0.10	n = 29	p = NS
Psychiatrist 4	r = +0.49	n = 34	p = 0.01
Pooled Psychiatrists	r = +0.19	n = 139	p = 0.01

and 11) were not used by the raters in this study. It is quite evident that Coder 1 and Coder 2, though exhibiting appreciable numbers of false positives, did yield significant effects in the expected direction (Tables 2 and 3). The most serious errors are to be found in the lower left and upper right cells. Coder 1 was not "embarrassed" by having any cases where, having predicted a low level of likelihood for future violence, individuals committed acts considered to be of high dangerousness by the external raters. But she did have 28 false positives against her. Whether these were "real" errors in prediction or were due to the fact that in all likelihood much violent behavior during follow-up goes unrecorded is a point we mention again below. The performance of Coder 2 was very similar to that of Coder 1. Though making three fewer definite false positive errors than Coder 1 she did, however, make two definite false negative decisions. It is clear that as well as a simple correlation between prediction and outcome it is necessary to examine the kinds of errors made. A false positive error has very different effects than a false negative error.[10] Table 4 shows the same general trend across the four pooled psychiatrists. Data are pooled because the number of patients assessed by

Table 2. Coder 1's Prediction of Future Dangerousness to Others Against Actual Outcome as Judged by Independent Raters

	Actual Outcome			
Prediction	**Low (1-3)**	**Med (4-5)**	**High (6-8)**	**Total**
Low (1-3)	12	13	0	25
Med (4)	10	10	6	26
High (5-7)	28	32	29	89
Total	50	55	35	140

$\chi^2 = 11.183$, df = 4, $p < 0.05$

Table 3. Coder 2's Prediction of Future Dangerousness to Others Against Actual Outcome as Judged by Independent Raters

	Actual Outcome			
Prediction	**Low (1-3)**	**Med (4-5)**	**High (6-8)**	**Total**
Low (1-3)	16	14	2	32
Med (4)	15	25	7	47
High (5-7)	25	23	25	73
Total	56	62	34	152

$\chi^2 = 14.825$, df = 4, $p < 0.01$

Table 4. Pooled Interviewing Psychiatrists' Predictions of Future Dangerousness to Others (N=4) Against Actual Outcome as Judged by Independent Raters

	Actual Outcome			
Prediction	**Low (1-3)**	**Med (4-5)**	**High (6-8)**	**Total**
Low (1-3)	11	7	6	24
Med (4)	15	17	7	39
High (5-7)	34	27	19	80
Total	60	51	32	143

$\chi^2 = 1.741$, df = 4, $p > 0.05$

each clinician was relatively small. Taken together the psychiatrists had a fair number of true positives to their credit (19) but this was offset by a comparatively large number of definite false negatives (6). Moreover, relative to the two external coders they had a large number of definite false positives (34). The overall result, shown in Table 4, was not significant.

Two final points from Tables 2, 3, and 4 merit note. The two external coders taken together considered roughly two-thirds to half of the persons to be highly dangerous at assessment. But at follow-up only about a quarter were actually detected to have committed acts rated by others as being of high dangerousness. This, it would seem, is in accord with the bulk of the results from the research literature on the prediction of violent behavior. Yet had the follow-up interval been longer than two years, some of those false positives might have become true positives. We are therefore anxious to repeat our follow-up search of records at five years post-assessment. The number of false positives can be expected to drop by some as-yet-unknown amount as a result of the former patients' increased opportunity to engage in violent conduct afforded by the mere passage of time. We touch on this point below.

The second point is that the psychiatrists in this study, though less accurate than the external coders in their predictive ability, did not attribute more dangerousness to the patients than did the external coders. This can be seen from the information in the right-hand column of Table 2, 3, and 4. In other words, the relative inferiority of the pooled psychiatrists was not simply due to the fact that, generally, they were especially apt to "find dangerousness" in their patients. Such a finding that psychiatrists are apt to "construct"[11] such dangerousness would not necessarily be surprising since, it is they, not external research coders, who carry actual responsibility in making recommendations to the court and they who must suffer the professional consequences for erroneous opinions. Other researchers have found that psychiatrists are more prone than members of other disciplines to impute such dangerousness.[12]

Discussion

We report two main findings. First, it would appear that two previously untrained coders can grasp essentially the same meaning of such a seemingly amorphous term as "dangerous to others in the future." That is, they can achieve acceptable levels of inter-rater reliability. To what extent this was facilitated by the fact that the coders were using "dangerousness to others in the future" as a global term at the end of a scale containing more specific items is not known. Second, predictions once made do relate to outcome. A third point, more-or-less incidental since we established it before, is that despite rather wide individual variations, some psychiatrists are able to predict future dangerous behavior to some degree. This limited predictive power apparently was not enhanced by knowledge that predictions would be checked against outcome. In this study the overall pooled correlation coefficient was +0.19 whereas in the previous one it was +0.20 (with individual clinicians showing considerable variation in ability between studies).

The first issue to be addressed is the relative consistency of the coder's predictions against the variability of the psychiatrists' predictions. With so few coders and psychiatrists it would be unwise to speculate too widely. Yet it seems to us probable that the two coders, coming to the study inexperienced, were influenced by the training scheme we imposed. As a result of this training they may have come to view the patients in something like the same light. In contrast, the psychiatrists, despite having been participants in many prestudy discussions about the meanings of terms in the DBRS, probably remained influenced by their prior training and experience. Although they may have attempted to adopt the investigators' perspective for the purposes of the study, it is likely that they were unable to approach their patients with a totally fresh outlook (or, indeed, that they could or ought to have done so insofar as they remained legally responsible for the assessments). It also is important to note that the present study was not so much a test of ability to predict dangerous behavior but of the coders' and psychiatrists' ability to "fit" the outlook created by researchers (though admittedly with much participation from clinical staff) and, indeed, the point of view of the criminology graduate students who scored the outcome measures. Considering the wide variability among forensic psychiatrists in how they view dangerousness,[13] and indeed among physicians in how they interpret terms such as "low probability" and "normally,"[14] it is hardly surprising that we achieved in this study such marked inter-psychiatrist differences. This much said, the pooled psychiatric correlations were moderately good. Rosen recently found that forensic psychiatrists show, as might be expected, wide variability in the way they interpret adjectives such as "probably" and "likely" as they modify statements about dangerousness.[15]

A critic might argue about the term "moderately" good.[16] What is a moderately good correlation? It must be conceded right away that correlations of the magnitude reported here account for but a small fraction of the total variance. Had we not had a fairly large sample we would not have been able to demonstrate statistical significance of the correlations. Even so, with limits of +1.0 and -1.0 a reader is entitled to ask why the strength of prediction-outcome correlations was not stronger and therefore more convincing. A coefficient of, say, +0.75 between psychiatric prediction and outcome would have been altogether more reassuring and could have been taken as much-called-for evidence that psychiatrists and other mental health workers can and perhaps should offer predictions about the possible future violent conduct of their patients.

We argue that though a prediction-outcome correlation of about +0.20 would be modest indeed for many tasks, it is, given the nature of the present undertaking, fairly good. Consider a professor who wishes to undertake a rather simple prediction exercise with a class of 20 or so students. If he or she correlates scores on a mid-term test with scores on the final examination, a coefficient of about +0.70 can be expected. Note that though this indicates a strong relationship, it is far short of perfect. Moreover the circumstances are more-or-less ideal in that students, unlike forensic psychiatric patients, will have good reason to disclose everything they know about the topic. All the students will show up for the exam, and the data will be complete and "clean." This can be contrasted with

the present kind of prediction-outcome task where the postassessment environments are varied and largely uncontrollable.

There are various reasons for the present correlations being low relative to simpler more-or-less ideal prediction-outcome tasks. Twelve of these are listed; the first 11 are dealt with in point form. An additional point, a particularly important one, is then covered in more detail.

Among the many reasons why the present pooled psychiatric and rater correlations do not on average rise much above about +0.20 are: (1) the clinical examination in this study was brief — it is possible that fuller explorations might have yielded more accurate predictions; (2) predictions of dangerous behavior, if transmitted to the court, may have influenced the judge to award prison incarceration or to seek hospital confinement for the offender — these very practical predictions made as a matter of routine may have reduced the opportunity for violent behavior to occur; (3) clinicians in this study were obliged to make predictions in all cases regardless of whether they had confidence in their ability to make accurate prognostications in specific cases — this forcing of responses likely introduces some degree of error; (4) outcome measures in studies of the present kind are hard to obtain and undoubtedly greatly underestimate the actual amount of violent behavior occurring in a follow-up period — at least some of the false-positive errors are not so much clinical mistakes but "research errors"; (5) follow-up intervals restricted to two years may allow insufficient time for the occurrence of particular violent behavior — with a long time available, the prediction has an increased chance of being fulfilled; (6) "dangerousness," though defined to some degree in the present study by the DBRS Scale, is nonetheless a complex construct — the clinicians cannot have had in mind a clear and commonly held idea of what kinds of events they were supposed to be predicting; (7) there is the point that it will have been decidedly to the advantage of some patients not to have disclosed information that would be pertinent to an assessment of future violent conduct — the evaluations will have been deliberately distorted by some patients; (8) some patients will receive inadequate or inaccurate assessments due to how they are perceived or regarded by individual clinicians — some of the error in such complex predictions would seem due to specific emotional, attitudinal, and other limitations of the individual clinicians; (9) certain kinds of information made available to clinicians at the outset of an assessment may direct the examination into paths unlikely to increase predictive accuracy — police reports and the like may induce clinicians to establish theories about individuals that stand in the way of accurate and unbiased assessment; (10) patients remanded for psychiatric assessments by the court appear for examination at what might be called "peak periods" of their lives — the circumstances they face at that time are to some extent atypical and unrepresentative of their ordinary life circumstances; and (11) correlations might be expected to rise slightly if the clinicians rather than uninvolved raters had scored the incidents obtained during the two-year follow-up period — there may have been mismatches between standards for dangerous conduct at time of assessment and at time of follow-up.

A further possible reason for our not having achieved higher prediction-out-

come correlations than were found is that there appears to be a ceiling for predicting future conduct from traits like honesty, conscientiousness, friendliness, and by extension, dangerousness. The implications of this have been dealt with by Mischel,[17] by Bem and Allen,[18] and by Mischel and Peake.[19] The general point is that when trying to make specific predictions about future behavior on the basis of information inferred from personality traits, it may never be possible to cross a "sound barrier" set around +0.40.[20] Clinicians tend to base their prognostications largely on the patient before them. Without full knowledge of the crucial situational factors that influence the occurrence or nonoccurrence of violent behavior it is likely that the clinician will be drawing from the "wrong" source.[21] It is tempting to infer personality traits especially as our language abounds with traitlike words and concepts; it is much harder to think in situational terms. As well, the clinician's very presence induces the patient under assessment not to show the full range of behavior of such critical interest. It is difficult to remember that just because some characteristics, such as intellectual ability and physical appearance, remain rather stable across a variety of situations, other behaviors can be expected to show no such cross-situational consistency. Clinicians, no less than other people, likely tend to overgeneralize the consistencies actually present. Instruments such as the DBRS mentioned here and discussed in more detail elsewhere[22] may be inadequate (and too apt to have low predictive ceilings) precisely because they place too great a reliance on certain qualities and traits expected to be present in and equally important to all individuals under assessment.

In other words, our "nomothetic" methods in which all individuals are scaled on the same dimensions, some of which could be utterly irrelevant to particular individuals, may face very large limitations when applied to exceedingly complex prediction problems where one or a few highly idiosyncratic circumstances may be exerting profound effects. If any of the above points hold true (and we think they do), it means that as researchers we may have to devise "idiographic" predictive methods that have at their center the particular patient's constructs and not the rather rigid and uniformly applied parameters of researchers. Of considerable interest is the idea that some patients may be more predictable than others in terms of characteristics while other patients need to be viewed largely against a background of environmental variables. It seems certain that efforts of the kind reported here, though well worth pursuing, suffer from serious inherent limitations of a kind not normally recognized in forensic psychiatry.

We suggest, given the complexity of the prediction task, there is reason to be encouraged by the present findings. Perhaps it is such data that recently have led Monahan to comment, "Empirically it is much less clear to me than it once was that relatively accurate prediction is impossible under all circumstances."[1] However, the present findings endorse our view that we pose a poor and misleading question when we ask, "Can psychiatrists predict dangerous behavior?" Results of the kind outlined here should encourage clinicians to ask other kinds of questions, all of which go beyond the present data. What cues do good predictors rely on as they form their opinions? Which kinds of patients are the most predictable? What specific kinds of future violent action is or is not likely to be expected from

which kinds of patients? How good are patients themselves at predicting their own future violent conduct? To what extent is our +0.20 correlation an *overestimate* of clinical acumen attributable, not to clinical opinion per se, but to the use of background information of the kind normally provided to psychiatrists in police reports? We think such questions can be answered in part through studies like the one described here (loose control of variables somewhat offset by relatively large numbers of subjects), through examination of short-range predictions where behavior can be monitored accurately in fairly controlled hospital-type environments,[23] and perhaps most interesting of all, through clinicians themselves becoming more research minded in establishing and checking specific predictions for and with their particular forensic psychiatric patients.

Acknowledgments

The authors are indebted to the clinical staff of the Brief Assessment Unit, Metropolitan Toronto Forensic Service, whose patient cooperation greatly facilitated the collection of data for this study. Follow-up of subjects was made possible by the diligent assistance offered by the medical records staff at each of the five psychiatric facilities cited and by the Research Division of the Ontario Ministry of Correctional Services.

Financial support for this study was provided by the Ontario Ministry of Health under Grant DM395, the Department of Justice, Canada under contract number 01GS 19200-8-0029, and the Centre of Criminology, University of Toronto.

References

1. Monahan J: The Clinical Prediction of Violent Behavior: An Assessment of Clinical Techniques. Beverly Hills, Sage, 1981
2. Steadman H and Cocozza J: Careers of the Criminally Insane. Lexington, MA, Lexington Books, 1974
3. Thornberry T and Jacoby J: The Criminally Insane: A Community Follow-up of Mentally Ill Offenders. Chicago, University of Chicago Press, 1979
4. Mullen JM and Reinehr RC: Predicting dangerousness of maximum security forensic mental patients. J Psychiat Law Summer:223-31, 1982
5. Sepejak DS, et al: Clinical predictions of dangerousness: Two-year follow-up of 408 pre-trial forensic cases. Bull Am Acad Psychiatry Law 11:171-81, 1983
6. Webster CD, Menzies RJ, and Jackson MA: Clinical Assessment Before Trial: Legal Issues and Mental Disorder. Toronto: Butterworths, 1982, see especially Appendix H
7. Slomen DJ, et al: The assessment of dangerous behaviour: Two new scales. Unpublished METFORS Working Paper in Forensic Psychiatry No 14, 1979
8. Megargee E: The prediction of dangerous behavior. Crim Just and Behav 3:3-21, 1976
9. It is of course always difficult to know how best to treat the cases where there is no information available at follow-up. They can be eliminated (as here), or they can be treated as zeros (since they were not involved in further incidents according to the record, they are best assumed to have been nondangerous). We have dealt with this topic at length elsewhere - see Menzies, Webster, and Sepejak, referred to in note 20 below.
10. Walker N: Dangerous people. In Weisstub DN (Ed.): Law and Psychiatry: Proceedings of an International Conference held at the Clarke Institute of Psychiatry, Toronto, Canada, February, 1977. New York, Pergamon, 1978
11. Pfohl SJ: Predicting Dangerousness: The Social Construction of Psychiatric Reality. Lexington, MA, Lexington, 1978
12. Quinsey VL: The long-term management of the mentally abnormal offender. In Hucker SJ, Webster CD, and Ben-Aron MH (Eds.): Mental Disorder and Criminal Responsibility. Toronto, Butterworths, see esp. 148-52
13. Menzies RJ, Webster CD, and Butler BT: Perceptions of dangerousness among forensic psychiatrists. Compr Psychiatry 22:387-96, 1981
14. Bryant GD and Norman GR: Expressions of probability: Words and numbers. N Engl J Med 302:411, 1980
15. Rosen J: How likely is likely? Complications arising from the use of words to estimate potential for dangerousness. Unpublished master's thesis in psychology, University of Toronto, 1983
16. We have recently been taken to task by some critics for describing such correlations as "moderately accurate" (see Stevens H and Bulmer R: Can Psychol 24:145-49). The explanation provided in the present text goes some way toward a defense of the terminology previously used by us.
17. Mischel W: Personality and Assessment. New York, Wiley, 1968

18. Bem D and Allen A: On predicting some of the people some of the time: The search for cross-situational consistencies in behaviors. Psychol Rev 81:506-20, 1974
19. Mischel W and Peake PK: Beyond déjà vu in the search for cross-situational consistency. Psychol Rev 89:730-55, 1982
20. Menzies RJ, Webster CD, and Sepejak DS: Hitting the forensic sound barrier: Predictions of dangerousness in a pre-trial clinic. In Webster CD, Ben-Aron MH, and Hucker SJ (Eds.): Dangerousness: Probability and Prediction, Psychiatry and Public Policy. New York, Cambridge University Press, in press
21. Felson RB and Steadman HJ: Situational factors in disputes leading to criminal violence. Crim 21:59-74, 1983
22. Sepejak DS, Webster CD, and Menzies RJ: The clinical prediction of dangerousness. Getting beyond the basic questions. In Müller DJ, Blackman DE, and Chapman AJ (Eds.): Psychology and Law: Topics from an International Conference. Chichester, UK, Wiley, in press
23. Rofman ES, Askinazi C, and Fant E: The prediction of dangerous behavior in emergency civil commitment. Am J Psych 137:1061-64, 1980 □