Brit. J. Psychiat. (1984), **144,** 149–155

The Reliability of Dangerousness Assessments
A Decision Making Exercise

CLÉOPÂTRE MONTANDON AND TIMOTHY HARDING

Summary: One hundred and ninety three raters drawn from six countries and representing different professional groups considered 16 case histories. Raters were asked to indicate what degree of 'dangerousness' they attributed to each individual and what they considered to be the optimal management.

The level of agreement between raters concerning the assessment of dangerousness was generally low, the level of 60 per cent being reached for only 4 cases out of 16. Psychiatrists did not reach a higher level of agreement on the ratings of dangerousness than non-psychiatrists. Psychiatrists had a tendency to rate individuals as more dangerous than did non-psychiatrists.

The results of this exercise do not support the use of 'dangerousness' as a scientifically or operationally valid concept.

The expertise of the psychiatrist in the assessment of dangerousness has been recognized as having special weight and authority (Stone, 1975). There is a feeling that in some way the psychiatrist knows better in matters of behaviour, with his training in observation and diagnosis, than people relying on their intuitions.

Nevertheless, a number of studies have indicated the questionable *validity* of predictions of dangerousness, as well as the many non-psychiatric elements which influence these predictions (Wenk *et al,* 1972; Frederick, 1978; Pfohl, 1978). However, the related questions of (a) the *reliability of the assessment of dangerousness,* and (b) the *relationship between assessments of dangerousness and decisions about patient management* have not received, to our knowledge, attention. The purpose of this study was to examine the degree of reliability or agreement between assessments of dangerousness and to compare decisions on dangerousness with decisions about management. Since most recent studies on dangerousness have been carried out in North America, we sought to widen the field of our enquiry. The study was therefore carried out in six countries with varying legal systems and socio-economic backgrounds.

Reliability studies test the precision of definition of the concept to be tested as well as the training and skill of the raters. A reliability study concerning, for example, the diagnosis of mental disorders will show whether raters use the same coding for the same patients. Unreliability may be due to the inadequacy or structural weakness of the classification system, or to imprecise definitions of the items in the classification or to limitations of the psychiatric knowledge of the raters (i.e., insufficient or inconsistent training). The validity and the reliability of psychiatric diagnosis in general have been considered insufficient by many studies (Zubin, 1967), and many researchers, some supported by WHO, have investigated the problem of variation in psychiatric diagnosis and their classification (Stengel, 1959; World Health Organization, 1975).

When we apply the concept of reliability (i.e, 'agreement between raters') to the assessment of dangerousness, contrasts are immediately apparent as compared to reliability in diagnosis (Shepherd *et al,* 1968; Astrup and Odegaard, 1970; Rutter *et al,* 1969; Averbuch *et al,* 1968; Fifth WHO seminar on psychiatric diagnosis, classification and statistics, 1972; Helmchen *et al,* 1973; Shepherd and Sartorius, 1974). 'Dangerousness' can be analysed in several ways: the person(s) or objects exposed to danger and the nature of the potential harm; the period over which a prediction of dangerousness is applied; and the way in which the label 'dangerous' is applied. A review of existing mental health legislation (Curran and Harding, 1978) and a semantic analysis of legal definitions of dangerousness (Harding, 1981) have shown that despite theoretically complex nature of the dangerousness criterion, it is generally used in a unitary fashion without explicit or clear definitions. We took as a starting point the observations that psychiatrists make assessments of dangerousness and set out to examine what level of agreement they achieve in such assessments.

https://doi.org/10.1192/bjp.144.2.149 Published online by Cambridge University Press

It seems clear that a certain degree of *reliability* in dangerousness assessments is a prior requisite for their validity: if we discovered high degrees of reliability among those with specialist training in questions of human behaviour, then we would be led to conclude that there *may* be some operational validity in their assessments. In order to compare the reliability of psychiatrists' assessments we included other medico-social professionals, magistrates, lawyers, police officers, as well as a certain number of persons who were not professionally involved in the evaluation of danger.

The study took place in six countries which were selected on account of their legal system, pattern of mental health services, socio-economic development, availability of research workers and desire to participate: Brazil (Province of Rio Grande do Sul), Denmark, Egypt, Swaziland, Switzerland (Canton of Geneva) and Thailand.

This research was part of a collaborative study (Soothill *et al*, 1981a; Soothill *et al*, 1981b) on the assessment of dangerousness carried out in six countries and coordinated by WHO as part of its medium-term programme in mental health. The study described here has already been repeated in Bulgaria, the USA (in Arizona and Massachusetts) and Canada: further comparative research findings may therefore be anticipated.

Research design

The collection of case histories

The main purpose of this study was to find out how different professional groups in different countries agree or disagree on rating dangerousness when given standard information. This standard information was our basic material and consisted of case histories.

The case histories concerned real people and covered the widest possible range of problems, chosen from all participating countries and representing the main settings in which dangerous assessments are made (e.g. mental health area, criminal and civil law area, prison system). Seven case histories, as diversified as possible both in terms of psychiatric history (including psychotic persons, 'psychopaths', drug addicts or alcoholics, neurotic persons) and in terms of criminal or delinquent behaviour (including violent offences, property offences, sexual offences, family violence, etc.) were collected from each of the six participating countries. The country teams were asked to give a standard format in the presentation of the histories (respecting anonymity, including information on sex, age, socio-economic status, mental status, criminal record, etc.). From a total of 42 case descriptions, 16 histories were finally selected for the reliability exercise. This number was a compromise between the ideal of presenting to the raters a diversity of cases and the need to limit the total number of cases to avoid overtaxing the time and motivation of the raters. A pilot test was carried out with the participation of the main investigators and technical advisers for the study.

There were four case descriptions with histories of violence associated with mental illness (including organic and functional psychoses, neurotic disorders, but *not* personality disorders); four case descriptions of individuals who had carried out violent acts, but were not mentally ill (although possibly suffering from a personality disorder); five case descriptions of persons mentally ill but non-violent; and three cases concerning individuals not mentally ill (but suffering from a personality disorder) with no history of violent acts. (In the text we have given the following codes to these four groups of case histories: V:MI = violent and mentally ill; V:NMI = violent and not mentally ill; NV:MI = non-violent and mentally ill; NV:NMI = non-violent and not mentally ill).

It was hypothesized that the cases belonging to the first and last categories, those who had a clear diagnosis of mental illness and presented violent behaviour and those who were not violent and not mentally ill, would be the object of high inter-rater agreement, particularly among psychiatrists.

Selection of participants for the 'reliability' testing

We planned that 35 raters from each country would participate in the reliability exercise, distributed in the following groups: 10 psychiatrists, 10 magistrates or persons involved in the criminal justice system (lawyers, jurists, policemen, prison officers and other legal officials), 5 persons belonging to medico-social professions and working closely with psychiatrists (psychiatric nurses, social workers, psychologists, general practitioners, etc.), 10 persons not involved in any medico-social profession and occupying some position of influence in the community (e.g. deputees, councillors, teachers, professors, village chiefs, religious leaders, unionists, etc.).

The final composition of the raters is shown in Table I. It differed from the ideal composition described above. The main discrepancies were that Egypt had a smaller number of raters than other countries and that psychiatrists were under-represented in Swaziland and over-represented in Egypt. The significance of these discrepancies will be considered in the discussion of the results.

The assessment questionnaires

Each of the 16 case descriptions led to a point when a decision had to be taken. Each participant was asked to put himself in the position of a person who had to make

https://doi.org/10.1192/bjp.144.2.149 Published online by Cambridge University Press

TABLE I

Raters by professional groups and by countries

	Brazil	Denmark	Egypt	Swaziland	Switzerland	Thailand	Total
Psychiatrists*	9	10	22	1	9	11	62
PJPs*	10	8	—	12	10	12	52
MSPs*	6	5	—	16	8	7	42
Others*	10	12	1	2	7	5	37
Total	35	35	23	31	34	35	193

*In this and subsequent tables we use the following code for professional groups.
1 Psychiatrists
2 PJPs = penal justice professionals (judges, lawyers, magistrates, police officers)
3 MSPs = medico-social professionals (doctors other than psychiatrists, nurses, social workers)
4 Others = people not belonging to the professions indicated above.

an assessment of the situation and reach a decision. Each case description was followed by a questionnaire.

The participants were asked to decide:

(a) What measure fitted most the person whose history they had just read (i.e. treatment in the community, admission to hospital, commitment to prison, etc.),
(b) The degree of dangerousness they attributed to this person—according to a four-point scale ranging from 'no dangerousness' to 'extreme dangerousness' (raters could also select a fifth item: 'impossible to estimate'),
(c) The certitude with which they made their assessment,
(d) Why they considered this person dangerous or not (open question),
(e) What factors influenced their assessment (from a given list),
(f) Whether the information given on the case was adequate.

The questions were closed, with the exception of question 4, but the participants could make additional 'open' comments for each question if they desired.

Results

Assessments of dangerousness

The central issue of the study concerned the degree of agreement among raters. We have quantified the level of agreement reached on ratings of dangerousness as the proportion of raters who agreed with the majority decision expressed as a percentage. Thus, if for a given case 110 of the 193 participants selected a rating of 'no dangerousness', this represented a level of agreement of 57 per cent. The question of an *acceptable* level of agreement is more of a value judgement than a statistical issue and may depend on the consequences for the individual being assessed—an analogy can be drawn to the degrees of proof required in legal proceedings (the degree of proof required in criminal cases involving serious consequences for the individual is very high, whereas it may be much lower in civil cases). An acceptable rate of agreement is therefore a matter of contention. For example, one may believe that in decisions which may have both serious and long-lasting consequences for an individual, a level of agreement above 90 per cent could be reasonably required. In decisions which may involve short-term and less serious consequences for a person, a level of agreement of 60 per cent may be considered reasonable. The majority ratings, the degree of reliability and the type of case history are shown for each case in Table II.

In 4 out of 16 case descriptions of inter-rater agreement in excess of 60 per cent was found. In cases 7, 10 and 14 there was no violence involved and the majority of raters (71.5 per cent, 85.5 per cent and 78.8 per cent respectively) indicated that there was no dangerousness. In case 15, which involved both violence and mental illness, dangerousness was rated as 'extreme' by 60.6 per cent of raters.

For four additional case histories (cases 1, 3, 11 and 12) a level of agreement of between 40 per cent and 60 per cent was recorded. The other eight cases present levels of agreement below 40 per cent. Considering these results by country we found the highest levels of agreement in Denmark (with agreement above 60 per cent in 6 cases, mainly with low degrees of dangerousness); and, in Thailand (agreement above 60 per cent in 4 cases, mainly with intermediate degrees of dangerousness).

For five of the case histories, more than 10 per cent of the raters indicated that dangerousness was 'impossible to estimate'; the overall rate for such ratings was 8 per cent.

In summary, of the 16 case descriptions subject to

https://doi.org/10.1192/bjp.144.2.149 Published online by Cambridge University Press

TABLE II

Cases indicating type, majority rating of dangerousness and level of agreement

Case number	Type of case*	Majority rating on 4 point dangerousness scale†	Level of agreement of dangerousness rating
1	V:MI	2	47.2
2	V:NMI	1	39.9
3	NV:MI	1	40.9
4	NV:MI	0	34.7
5	V:NMI	2	33.7
6	V:NMI	1	39.9
7	NV:MI	0	71.5
8	NV:NMI	0	30.1
9	V:MI	2	32.1
10	NV:NMI	0	85.5
11	NV:MI	0	46.6
12	V:NMI	2	47.7
13	V:MI	2	29.0
14	NV:NMI	0	78.8
15	V:MI	3	60.6
16	NV:MI	1	33.7

*V:MI = violent and mentally ill; V:NMI = violent and not mentally ill; NV:MI = non-violent and mentally ill; NV:NMI = non-violent and not mentally ill

†0 = no dangerousness; 3 = extreme dangerousness

ratings by 193 raters, a quite high agreement (i.e. more than 60 per cent) was reached on one quarter of the cases; a tendency to agree was observed in a further quarter; while for half the cases the overall level of agreement was little better than that expected by chance.

Ratings by psychiatrists compared to those by non-psychiatrists

One of our initial hypotheses was that psychiatrists, as compared with lay people and other professionals, would present higher agreement as a group, as a result of their specialized training and their more objective rating behaviour.

When we compared the levels of agreement achieved by psychiatrists with those of all other groups combined (i.e. 'non-psychiatrists'), there was no significant difference (computed by two-way analysis of variance (ANOVA) using arcsin transformation). On the individual case histories, there were two cases (11 and 14) in which the non-psychiatrists as a group displayed a significantly higher level of agreement ($P < 0.05$ for each case) than the psychiatrists. On one case (case 15) the psychiatrists had a higher level of agreement ($P < 0.05$).

It could be argued that our failure to demonstrate a higher level of agreement on *dangerousness ratings* by the psychiatrists as compared with non-psychiatrists stems from an artefact of experimental design. We can, however, refer to the results concerning *management decision*. Here, using the same methodological approach and statistical analysis, we did indeed demonstrate a statistically significant tendency for psychiatrists to agree more than non-psychiatrists ($P < 0.05$) on all cases combined. There were three individual cases (cases 9, 11 and 15) where the psychiatrists' level of agreement was significantly higher than that of the non-psychiatrists. In no case did non-psychiatrists reach a higher level on *management decisions* than psychiatrists.

We can summarize this central part of our findings as follows:

(a) Psychiatrists reach a higher level of agreement when asked to decide about the management of cases than non-psychiatrists, a finding which would be predicted given that psychiatrists are trained to make such decisions.

(b) This was not the case when we examined the level of agreement on dangerousness ratings, suggesting that a professional training in psychiatry does not lead to increased reliability of such ratings.

(c) A comparison between psychiatrists' ratings on dangerousness and on management decisions showed signigificantly higher levels of agreement on the latter. This finding lends further support to the suggestion that a psychiatric training does not necessarily have a direct relationship with increased reliability in dangerousness assessments, whereas it does favour agreement about the way in which patients should be managed.

Comparisons between professional groups

In order to examine the distribution of the dangerousness ratings by professions we present these ratings according to an arbitrary three-fold division of the average ratings of each participant. The 'average dangerousness rating' for each participant was calculated by adding up the numerical value given to the ratings made by each rater of all 16 cases, excluding the 'impossible to estimate' and the 'no answer' and dividing by the number of cases thus included.

Table III shows significant inter-professional differences. Psychiatrists give higher ratings of dangerousness as compared to the other groups (40.3 per cent being in the high rater group), followed by the medico-social professionals (38.1 per cent). The non-specialists (other) have the lowest tendency to give high ratings of dangerousness (24.3 per cent) while penal justice professionals are more evenly distributed.

Assessments of dangerousness by types of case history

It was hypothesized that inter-rater agreement

https://doi.org/10.1192/bjp.144.2.149 Published online by Cambridge University Press

TABLE III

Dangerousness ratings by profession (percentage)

	Profession				
Average rating	Psychiatrists %	PJPs %	MSPs %	Other %	Combined %
Low raters	17.7	38.5	40.5	40.5	32.6
Medium raters	41.9	32.7	21.4	35.1	33.7
High raters	40.3	28.8	38.1	24.3	33.7
Total	100	100	100	100	100
	(n = 62)	(n = 52)	(n = 42)	(n = 37)	(n = 193)

$\chi^2 = 11.72$ df = 6 $P < 0.05$

TABLE IV

Dangerousness assessment by type of case (percentages)

	Danger rating						
Type of case	None	Minor	Moderate	Extreme	Impossible to estimate	No answer	Total
MI + V	11.0	17.4	32.6**	25.6*	13.0	0.4	100
NMI + V	21.0	30.3**	29.0*	12.2	6.9	0.6	100
MI + NV	41.7**	29.4*	17.8	4.0	6.9	0.1	100
NMI + NV	64.8**	15.5*	10.0	4.5	4.7	0.5	100
	(n = 1024)	(n = 724)	(n = 706)	(n = 357)	(n = 247)	(n = 12)	(n = 3088)

**commonest rating *second commonest rating

would be higher when mental illness and violent behaviour were both present or when neither violent behaviour nor mental illness were present than for those cases where mental illness was present without violent behaviour or vice versa,

Table IV shows that, as would be expected, when violent behaviour was present, and particularly when it was associated with mental illness, dangerousness tended to be rated as high. Ratings were also high when violent behaviour was not associated with overt mental illness. When there was no violent behaviour, ratings were lower, and this was more marked when mental illness was also absent. The presence or absence of violent behaviour had a stronger effect on dangerousness ratings than the presence or absence of overt mental illness.

Management decisions

Treatment and/or supervision in the community was the most frequently recommended measure (50.8 per cent), followed by admission to hospital (32.8 per cent), by life in the community without any treatment or supervision (10.5 per cent) and by prison (5 per cent). Raters from Brazil and Swaziland advised treatment in the community more frequently than raters from other countries. Those from Denmark and Switzerland proposed re-integration in the community without treatment or supervision more frequently than raters from other countries. Raters from Egypt and Thailand recommended closed institutions more frequently than others.

The results were less differentiated according to professional groups. The psychiatrists were the least likely to opt for life in the community without treatment or supervision and recommended admission to hospital more often than the others.

As in the case of the assessment of dangerousness, measures proposed were related to two variables in the case histories: the existence of violence and the existence of mental illness.

Table V includes all raters. Mental illness appeared to be the first determining variable while violent behaviour is the second as far as management decisions are concerned.

We also compared the degree of reliability on management decisions with assessments of dangerous-

https://doi.org/10.1192/bjp.144.2.149 Published online by Cambridge University Press

TABLE V

Decisions according to groups of cases (percentages)

Group	Decision					
	In community			In institutions		Other response
	No treatment		With treatment	Hospital	Prison	
MI + V	1.8	45.8*	44.0	46.8	6.1	1.3
				52.9*		
MI + NV	7.5	52.9	45.4	45.7	0.7	0.7
				46.4		
NMI + V	6.6	70.3	63.7	19.4	9.5	0.8
				28.9		
NMI + NV	32.5	84.3	51.8	10.7	4.8	0.2
				15.5		
Combined	10.5	61.3	50.8	32.8	5.0	0.8
				37.8		

*combined total of two columns

ness. In 10 cases there was a significant difference above the P <0.01 level (i.e. no overlap of the 99 per cent confidence limits for the respective values)—in 6 of these there was greater agreement on *management decisions* than on *dangerousness assessments*. In general therefore there is a greater tendency to agree on *management* than on the more abstract concept of *dangerousness*. This tendency was more marked among psychiatrists than non-psychiatrists.

Reasons indicated for the assessment of dangerousness

Raters were invited to select from a list of seven factors the two factors which influenced their assessment of dangerousness the most.

The factors in order of frequency were: mental state, personality, personal history, social environment, nature of offence, criminal record. Raters who made high estimates of dangerousness tended to rely on 'mental state' and 'criminal record' as factors supporting their estimates; raters who made lower estimates tended to use 'social environment' and 'personality background' as determinant factors.

Discussion

We would summarize our main findings as follows. Psychiatists as a group tended to give the highest (most severe) ratings of dangerousness; the group of non-professionals gave the lowest ratings. The psychiatrists did not reach higher inter-rater agreement as regards the degree of dangerousness involved in the various cases than the other groups of raters. Psychiatrists were the least likely to indicate no danger, even for cases with no mental illness and no violent behaviour. Psychiatrists had a significantly higher level of agreement on management decisions than other groups, in contrast to their agreement on assessments of dangerousness. In only 4 cases (out of 16) did global inter-rater agreement on the degree of dangerousness exceed 60 per cent. Inter-rater agreement was highest for non-violent, not mentally ill individuals, who were assessed as being non-dangerous.

Several methodological and theoretical problems must be taken into account in interpreting these results.

As regards the selection of raters there is no doubt that the discrepancy between the ideal and the actual composition of the participants made analysis more complicated. The fact that the test was completed individually and with no time limit in some countries or collectively with a definite time limit in others introduced additional problems. The fact that some participants worked on translated versions (in French or Portuguese) added a further source of variation in results, while the use of the English version in non-English speaking countries introduced a bias in the selection of participants and possible difficulties in comprehension.

In order to have a more complete and comprehensive picture of what people in various countries mean by dangerousness, it would have been necessary on the one hand to present case descriptions to a wider sample of persons (our 'lay' people were generally privileged educationally and economically), and on the other hand to present not only dangerous acts due to particular mentally disturbed individuals, but also dangerous practices causing public or private harm, whether they be the products of 'normal' individuals or of corporate groups. In this respect it is significant that case 5

https://doi.org/10.1192/bjp.144.2.149 Published online by Cambridge University Press

presenting intra-family violence, which is not behaviour with high visibility and high reportability, was one of the cases with the lowest inter-rater agreement.

This research on assessments of dangerousness should be considered as a pilot study and the results tentative. The method seems to us to offer possibilities for examining the question of *reliability* of dangerousness assessment with larger samples of raters, without the added complication of transcultural comparisons. The insights of the participants, the problems that arose because of the collaborative nature of the reasearch and the experience gained can be used to refine the instrument, control more precisely the composition of the participants and broaden the scope of the cases described. The results of this study, with due regard to their limitations, do give some cause for concern: the concept of dangerousness is far from eliciting reliable responses, even from expert professionals.

Acknowledgements

This article is a report from the WHO Collaborative Study on Assessment of Dangerousness in Forensic and Administrative Psychiatry. Resources and technical support have been provided by the World Health Organization and national authorities concerned. The research was also supported by a contract (278-77-0015(SM)) from the National Institute of Mental Health in the USA. We wish to thank the country coordinators (Dr H. Adserballe, Denmark; Professor J. Bernheim, Switzerland; Dr T. Dasananjali, Thailand; Dr R. Rinto Ribeiro, Brazil; Dr F. Reinhold, Swaziland; Professor M. I. Soueif, Egypt) and other members of the country teams who participated in this project. In addition, we are grateful to Trung Ngo Khac who painstakingly coded all the material and to Sandra Doyle who facilitated the information flow.

References

ASTRUP, C. & ODEGAARD, O. (1970) Continued experiments in psychiatric diagnosis. *Acta Psychiatrica Scandinavica*, **46**, 180–209.

AVERBUCH, E. S., MALNIK, E. M., SEREBRJAKOVA, Z. N. *et al* (1968) *Mental Disorders of Old Age. Diagnosis and Classification of Psychiatric Disease in Old Age*. Leningrad: Institute of Psychiatry.

CURRAN, W. J. & HARDING, T. W. (1978) *Law and Mental Health: Harmonizing Objectives*. Geneva: WHO.

FIFTH WHO SEMINAR ON PSYCHIATRIC DIAGNOSIS, CLASSIFICATION AND STATISTICS Mental retardation. (1972) *American Journal of Psychiatry*, Supplement to **2128**, No. 11, 3–14.

FREDERICK, C. J. (ed.) (1978) *Dangerous Behaviour: A Problem in Law and Mental Health*. Washington, D.C.: United States Department of Health, Education and Welfare, pp. 153–85.

HARDING, T. W. (1981) Du danger, de la dangerosité et de l'usage médical de termes affectivement chargés. *Déviance et Société*, **4 (3)**, 331–48.

HELMCHEN, H., KIELHOLZ, P., BROOKE, E. M. & SARTORIUS, N. (1973) The classification of neurotic and psychosomatic disorders (Zur klassifikation neurotischer und psychosomatischer Störungen), *Nervenartz*, **44**, 292–99.

PFOHL, S. (1978) *Predicting Dangerousness*. Lexington, Massachusetts: Lexington Books.

RUTTER, M., LEBOVICI, S., EISENBERG, L., *et al* (1969) Psychiatric disorders of children: a tri-axial classification of mental disorders in children. An international study. *Journal of Child Psychology and Psychiatry*, **10**, 41–61.

SHEPHERD, M., BROOKE, E. M., COOPER, J. E., *et al* (1968) An experimental approach to psychiatric diagnosis. *Acta Psychiatrica Scandinavica*, **44**, Suppl. 201.

—— & SARTORIUS, N. (1974) Personality disorders and the International Classification of Diseases. *Psychological Medicine*, **4**, 141–46.

SOOTHILL, K. L., ADSERBALLE, H., BERNHEIM, J. *et al* (1981a) Social control of deviants in six countries. *Medicine, Science and the Law*, **21**(3), 31–40.

—— HARDING, T. W., ADSERBALLE, H. *et al* (1981b) Compulsory admissions to mental hospitals in six countries. *International Journal of Law and Psychiatry*, **4**, 327–44.

STENGEL, E. (1959) Classification of mental disorders. *Bulletin of the World Health Organization*, **21**, 601–63.

STONE, A. A. (1975) *Mental Health and Law: A system in Transition*. Rockville, Maryland: National Institute of Mental Health.

WENK, E., ROBINSON, J. & SMITH, G. (1972) Can violence be predicted? *Crime and Delinquency*, **18**, 393–402.

WORLD HEALTH ORGANIZATION (1975) *Schizophrenia: A Multinational Study—A summary of the Initial Evaluation phase of the International Pilot Study of Schizophrenia*. Geneva: Public Health Paper No. 63.

ZUBIN, J. (1967) Classification of the behavior disorders. *Annual Review of Psychology*, **18**, 373–406.

Cléopâtre Montandon, Ph.D.

Timothy Harding, M.B., M.R.C.Psych

Institute of Legal Medicine, University of Geneva, 8, Passage de la Radio, 1211 Geneva 4, Switzerland

Reprint requests: Dr N. Sartorius, Director of the Division of Mental Health, World Health Organization, 1211 Geneva 27, Switzerland

(Received 18 April; revised 2 June 1983)

https://doi.org/10.1192/bjp.144.2.149 Published online by Cambridge University Press