Brit. J. Psychiat. (1971), 118, 629–40

Differences in Usage of Diagnostic Labels amongst Psychiatrists in the British Isles*

By J. R. M. COPELAND, J. E. COOPER, R. E. KENDELL and A. J. GOURLAY

Introduction

This paper examines the differences in diagnostic practice and rating behaviour between groups of psychiatrists trained in different centres in the British Isles, using their observations on three video-taped psychiatric interviews. The findings are also briefly compared with similar data obtained in the United States.

These studies are complementary to others concerned with the diagnosis of patients admitted to State Hospitals in the United States and United Kingdom undertaken by the same U.S.–U.K. Diagnostic Project (Cooper *et al.*, 1969, 1970; Gurland *et al.*, 1969, 1970). These hospital admission studies showed that psychiatrists in New York State hospitals made many more diagnoses of schizophrenia than did their British counterparts in mental hospitals serving London. To collect a final diagnostic label, however, is to examine only a limited aspect of the diagnostic process, and a series of video-tape studies was organized so as to have available some more detailed information about the relationship between the rating of symptoms and abnormal behaviour and the resulting diagnosis.

The three patients in this paper were chosen because it was thought likely that they would produce interesting and important agreements and disagreements between American and British psychiatrists. It was expected that for Mrs. A, American psychiatrists would make more diagnoses of schizophrenia than the British; for Mr. B, that all would agree on a diagnosis of schizophrenia; and for Mr. F, that American psychiatrists would make a diagnosis of schizophrenia, in contrast to one of personality disorder by British psychiatrists. Although these three patients were chosen primarily for their international interest, it is clearly well worthwhile examining the results for differences between various groups of psychiatrists in the British Isles, since some 200 psychiatrists participated.

Previous studies in this field by other workers have concentrated on international differences, and have used comparatively small numbers of psychiatrists. Sandifer *et al.* (1968) used 30 filmed interviews showing each film to audiences of up to ten psychiatrists in North Carolina, London and Glasgow. They found that British psychiatrists tended to diagnose manic-depressive illness more readily than American psychiatrists, and *vice versa* with depressive neurosis. Katz *et al.* (1969), using one filmed interview, found that approximately one-third of the American audience diagnosed the patient as having some form of schizophrenia, one-third as being neurotic, and the remaining third as having a personality disorder. When the same film was shown to a British audience at the Maudsley Hospital, London, no diagnosis of schizophrenia was made, but 59 per cent of the British psychiatrists made a diagnosis of personality disorder, a further 22 per cent diagnosed psychoneurosis, 16 per cent suggested other diagnoses, and 3 per cent (one psychiatrist) made a diagnosis of manic-depressive psychosis.

More recently, Kendell and Sharpe have described some other results of the video-tape studies carried out by the U.S.–U.K. Diagnostic Project (Kendell *et al.*, 1971) (including the data in this paper), in which four video-taped interviews were used to demonstrate how psychiatrists in the United States and British Isles differ in their ratings and diagnoses. (Some of the results of

* (U.S.–U.K. Diagnostic Project)

these studies are mentioned briefly in this paper). The outstanding findings were that psychiatrists in the U.S.A. perceive many more symptoms than their British counterparts, and are much more likely to make a diagnosis of schizophrenia.

Method

Audiences were assembled in the following centres of the United Kingdom and the Republic of Ireland: Edinburgh, Glasgow, Manchester, Birmingham, Belfast, Dublin and London. This was done with the assistance of the Royal Medico-Psychological Association and the University Departments in the chosen centres. At each centre a one-day meeting was held at which three video-tapes were rated.

In all, some 200 psychiatrists participated; all had had at least four years full-time psychiatric experience and possessed a Diploma in Psychological Medicine. The participating psychiatrists were volunteers, but care was taken to have a large proportion from mental hospitals, since these are the largest group and the ones who carry out the bulk of psychiatric work. Each video-tape lasted between 30 and 40 minutes and consisted of an unstructured clinical interview intended to elicit the principal symptoms of the present mental state. These interviews were not intended to be complete case presentations, and no additional written history was supplied.

Ratings were made on each tape, using the Lorr In-Patient Multi-dimensional Psychiatric Scale (IMPS) (Lorr *et al.*, 1962). This scale comprises a series of 89 ratings of abnormal behaviour in which each item is briefly defined in non-technical terms. One of the advantages of the IMPS is the comparatively simple and rapid method of analysis developed by Lorr, in which the ratings are summarized as a profile of standard scores on ten 'syndromes' derived by factor analysis. These syndromes do not correspond exactly with conventionally accepted clinical symptom groups, but they are similar enough for the scores on them to be used for the detection of major rating differences, and for the detection of consistent differences in the levels of rating of what is normal or abnormal.

Each syndrome represents a group of symptoms. (1) 'excitement' (EXC), covers hurried, loud speech that is difficult to stop, elevated mood and histrionic behaviour, (2) 'hostility' (HOS) covers observed hostility, (3) 'paranoid' (PAR) covers ideas and delusions of persecution, (4) 'grandiosity' (GRN) includes boasting, unrealistic optimism and ideas and delusions of grandeur, (5) 'perceptual distortion' (PCP) includes hallucinations, (6) 'introunpunitiveness' (INP) covers mainly depressive symptoms, particularly with guilt and self-depreciation, (7) 'retardation' (RTD) covers slowness of movement and speech, (8) 'cognitive disorientation' (DIS) includes disorientation in time and place, (9) 'motor' (MTR) includes abnormal or excessive movements and general activity, (10) 'conceptual disorganization' (CNP) includes disturbances of thought and language.

For comparing the syndrome scores derived from the ratings the method used was that of Greenhouse and Geisser (1959). This tests whether the profiles of the group mean syndrome scores are (i) on the same level, and (ii) have the same shape.

One of the main interests was to see whether there were any differences in the profiles in these two respects for the various centres of training, so for many of the analyses the ratings for each patient were divided into groups according to the centre where the rater had spent his first four years of full-time psychiatry, i.e. Maudsley Hospital, Other London Hospitals, English Provinces, Edinburgh, Glasgow, Dublin, Belfast and 'Others'. Personal information, such as age, type and length of psychiatric experience, was also collected from each rater; these details are summarized in Table I.

The raters were also asked to make a diagnosis according to the Eighth Edition of the International Classification of Diseases and to add to that a subsidiary and alternative diagnosis if necessary. Since the video-tapes were not complete case presentations, the raters were asked to indicate their confidence in their diagnosis.

Results

The differences in the diagnoses and IMPS ratings were examined by dividing the psychiatrists into groups by age, place and type of training, and type of main experience, to see which of these rater variables contributed the most. Differences for the diagnoses are presented here first, followed by those for the IMPS syndromes.

(a) *Diagnoses for the three video-tapes*

Table II shows the diagnoses made by the British and Irish centres on Mrs. A. The term 'affective psychosis' includes both depressive and manic types. (At the bottom of the Table are given the overall U.S. and British Isles findings: the American psychiatrists were all Board-eligible (see Kendell *et al.*, 1971).

https://doi.org/10.1192/bjp.118.547.629 Published online by Cambridge University Press

TABLE I

Personal details of total number of raters (n = 222)

		n	%
First four years of training at:	Area mental hospital	86	39
	University dept. of psychiatry	82	37
	Psychiatric dept. of medical school	27	12
	Other	27	12
Main Psychiatric Experience at:	Area mental hospital	83	37
	Psychiatric dept. of general or teaching hospital	47	21
	University dept. of psychiatry	57	26
	Other	35	16

Age			Years of experience			Psychotherapeutic training		
	n	%		n	%		n	%
Under 35	40	18	4–9	99	45	None	171	77
36–49	147	66	10–19	87	39	Freudian or Neo-Freudian	24	11
Over 50	35	16	Over 20	36	16	Other	27	12

TABLE II

Diagnoses made on Mrs. A by psychiatrists trained at different centres. Percentage table. (**indicates* $p < 0{\cdot}05$ *for difference*)

Training centre	Main Diagnosis: Affective psychosis	Schizophrenia	Other	per cent	Total
Maudsley	23	75	2	100	n = 44
Other London	21	75	4	100	n = 44
English provinces	20	77	3	100	n = 30
Edinburgh	26	69	5	100	n = 19
Glasgow*	58	42	—	100	n = 12*
Dublin	36	64	—	100	n = 14
Belfast	29	71	—	100	n = 14
Other	35	65	—	100	n = 34
Total British Isles	28	70	2	100	n = 211
Total U.S.A.	7	89	4	100	n = 122

Mrs. A was a woman with many severe symptoms that could be described as both schizophrenic and affective; in her case the affective symptoms were suggestive of a mixed manic-depressive type. She said that she felt wonderfully happy, but that something was forcing her to dance and sing and move her left leg. Messages were passed to her over the television, and at times people laughed at her because she must be in hell. She felt 'dirty and nasty' and thought she might have died as a child. She gave a history of a previous out-patient and in-patient treatment for depression.

A reasonably good agreement was obtained between the centres, with the exception of Glasgow which tended to favour a diagnosis of manic type of manic-depressive psychosis (58 per cent) rather than a diagnosis of schizo-

phrenia (42 per cent). Although numbers are small when Glasgow is compared with the other centres together, the difference is significant at the 5 per cent level. On the other hand, 89 per cent of American audiences considered this patient to be suffering from schizophrenia and only 7 per cent as suffering from an affective psychosis, so that Glasgow is most un-American in this respect. This is in keeping with the previous results found by Sandifer *et al.* (1968).

Raters whose main experience was in University Departments diagnosed Mrs. A as having an affective illness significantly more frequently than those whose main experience was in area mental hospitals ($p < 0.5$). In none of these three patients were there any correlations between diagnosis and such personal characteristics of the raters as age, length of time qualified, psychoanalytic or other type of training, or where the first medical qualification was obtained.

Table III shows the diagnoses made on the second patient. Mr. B was a man of fairly well preserved personality and affect who had suffered persecutory voices in clear consciousness intermittently for eight years. They commented upon what he was thinking, and he feared they came from a gang out to kill him. He had once thought that he had a radio-transmitter in his lung and that he was radio-active, but he now realized this was absurd. He had also been a consistently heavy drinker, but had stopped drinking seven months before the exacerbation of the illness during which he was interviewed.

For this patient all centres tended to agree on a diagnosis of schizophrenia and there is no significant difference between them. Edinburgh is more prepared than the others to consider alcoholic hallucinosis as an alternative diagnosis, but numbers are not large enough for the difference to reach significance. It is interesting that neither Dublin nor Belfast considered this diagnosis at all. This tape also produced the closest diagnostic agreement between the British Isles and the U.S.

When Maudsley raters are compared with Edinburgh raters alone, 95 per cent as against 70 per cent for Edinburgh diagnosed Mr. B as having schizophrenia (this difference reaches the 5 per cent level of significance).

Table IV shows the diagnoses made on the third video-tape. Mr. F was an unmarried American who had never held a steady job. He had a history of intermittent drug and alcohol abuse, and a brief paralysis of one arm which the patient said his doctor had called hysteria. He had undergone an unsuccessful course of psychotherapy. He also described how he had spent a lot of time lying in bed, drinking wine occasionally and watching television. He complained that his face often swelled and that his appearance changed. He gave graphic and colourful descriptions of his feelings, including his anxiety at the thought of work, and concern over his inability to keep friends.

Diagnostic agreement for this tape is high and no significant differences are apparent in the British Isles. Seventy-two per cent of diagnoses included under 'personality disorder'

TABLE III

Diagnoses made on Mr. B by psychiatrists trained at different centres. Percentage table

Training centre	Main Diagnosis				
	Schizophrenia	Alcoholic psychosis	Other	Per cent	Total
Maudsley	95	5	—	100	n = 41
Other London	95	3	2	100	n = 43
English provinces ..	90	10	—	100	n = 29
Edinburgh	70	18	12	100	n = 17
Glasgow	92	8	—	100	n = 13
Dublin	93	—	7	100	n = 15
Belfast	86	—	14	100	n = 14
Other	95	3	2	100	n = 57
Total British Isles ..	91	6	3	100	n = 205
Total U.S.A.	88	6	6	100	n = 17

https://doi.org/10.1192/bjp.118.547.629 Published online by Cambridge University Press

TABLE IV

Diagnoses made on Mr. F by psychiatrists trained at different centres. Percentage table

Training centre	Main Diagnosis: Affective psychosis	Schizophrenia	Personality disorder	Neurotic illness (mostly hysterical)	Per cent	Total
Maudsley	8	2	77	13	100	n = 39
Other London ..	8	2	63	27	100	n = 41
English provinces	—	—	75	25	100	n = 28
Edinburgh ..	—	—	81	19	100	n = 16
Glasgow	—	—	80	20	100	n = 10
Dublin	—	—	64	36	100	n = 14
Belfast	7	7	80	6	100	n = 15
Other	3	3	87	7	100	n = 31
Total British Isles..	4	2	75	19	100	n = 194
Total U.S.A. ..	8	69	7	14	100	n = 134

were specified as 'hysterical personality' and 72 per cent of those under 'neurosis' were specified as 'hysterical neurosis' so there is a high degree of agreement on some use of the term 'hysterical'. This is of particular interest because it was this tape which produced the widest disagreement between the British Isles and the U.S., where 69 per cent of raters made a diagnosis of schizophrenia. The commonest diagnoses in the U.S.A. were schizo-affective psychosis and paranoid schizophrenia, although 22 American psychiatrists said that they preferred to use 'pseudo-neurotic schizophrenia' (a term not recognized in the International Classification).

The significance of these findings must be seen against the psychiatrists' rating of confidence in their diagnosis. If the psychiatrist had considered that insufficient information was available on each tape to make a confident diagnosis the results would have little meaning. In fact, for Mrs. A over 70 per cent of psychiatrists at each centre were either 'quite confident' in their diagnosis, or 'reasonably confident' (rating of 1 and 2 on a scale of 1–5), and in four out of the seven centres well over 80 per cent indicated this. For Mr. B and Mr. F over 75 per cent at each centre and over 80 per cent at six centres indicated similar degrees of confidence.

(b) *Profiles from the IMPS ratings*

The IMPS syndrome profiles derived from the ratings on Mrs. A are shown in Figure 1. (There were 211 raters in the British Isles and 122 raters in the U.S.A.) The shaded area shows the spread of the mean scores of the different training centres in the U.K. The shapes of the profiles of syndrome scores of the U.K. raters are not significantly different from each other, but within this overall similarity of pattern there are significant differences on some of the individual syndromes. The most striking instance is Glasgow, which scores higher than the rest on syndromes 1, excitement, and 10, conceptual disorganization. These high scores are presumably related to another feature of raters trained in Glasgow, a preference for a diagnosis of mania for this patient. The profile of the American raters is fairly consistently above that of the British Isles raters.

Figure 2 represents the profiles for Mr. B. (There were 205 raters in the British Isles but only 17 raters in the U.S.A.) The profile shapes here are even more similar than those for Mrs. A, but even within this fairly narrow range of variation (shown on Fig. 2 by the shaded area) there are some significant differences between some of the centres on the individual syndrome means, in that the means of the Maudsley-trained raters are consistently

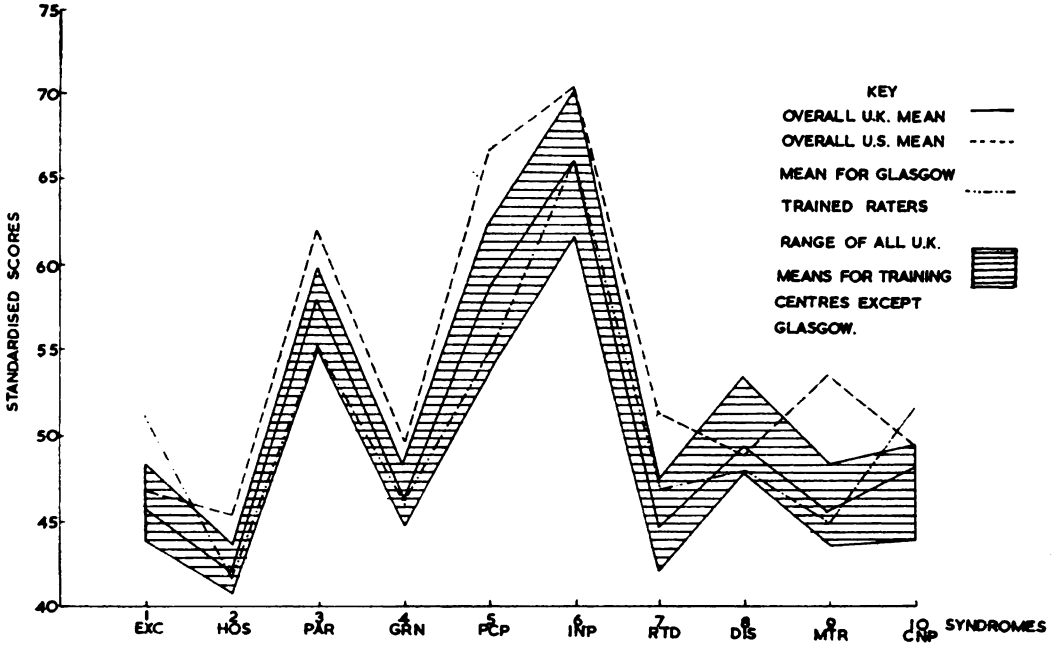

FIG. 1.—Inpatient multidimensional psychiatric scale. Mean symptom profiles of patient Mrs. A for raters from U.K. centres of training.

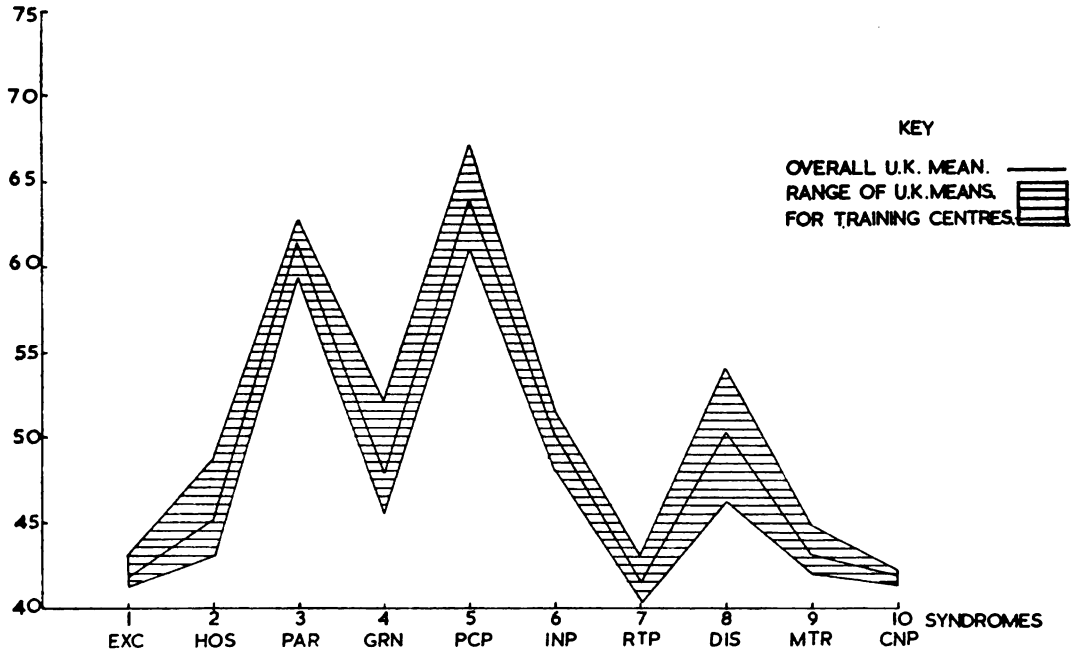

FIG. 2.—Inpatient multidimensional psychiatric scale. Mean symptom profiles of patient Mr. B for raters from U.K. centres of training.

https://doi.org/10.1192/bjp.118.547.629 Published online by Cambridge University Press

low, and those of the Glasgow raters consistently high.

Figure 3 shows the profiles for Mr. F, on whom there was most diagnostic agreement within the British Isles and least between the British Isles and the U.S.A. (There were 194 raters in the British Isles and 134 raters in the U.S.A.) Again the profiles from British Isles centres are more or less parallel, but for this tape the differences between the group mean scores of the syndromes are not significant. The raters in the U.S.A. once again rated symptoms on the whole more severely than the raters in the British Isles. Thus, in contrast to the dissimilarity of the diagnoses between the U.S.A. and the British Isles for this patient, there is a marked similarity of symptom pattern. Psychiatrists in both countries seem to perceive the same groups of symptoms, but they interpret them in a different way. It was thought possible that the severity of the symptoms, rather than their pattern, could be concerned in this American tendency to diagnose schizophrenia (Kendell *et al.*, 1971), but for both the American and the British raters there was no relationship between diagnoses and different levels of rating.

The effect of personal characteristics of the raters

Age did not have any relationship with diagnostic habits for these video-tapes. Examination of type of work against diagnosis produced only one significant finding, in that the psychiatrists whose main experience was in mental hospitals were more likely to make a diagnosis of schizophrenia for Mrs. A than those whose main experience was in University Departments of Psychiatry. This does not seem to be simply the effect of training, since a similar examination of diagnosis against place of the first four years of experience in psychiatry showed only a trend of the same type that did not reach significance.

For the IMPS syndrome scores, however, a number of significant effects emerged for place of training, age, and type of work. The most consistent finding was that raters trained at the Maudsley rated at a lower level than the rest. Taking all three tapes together

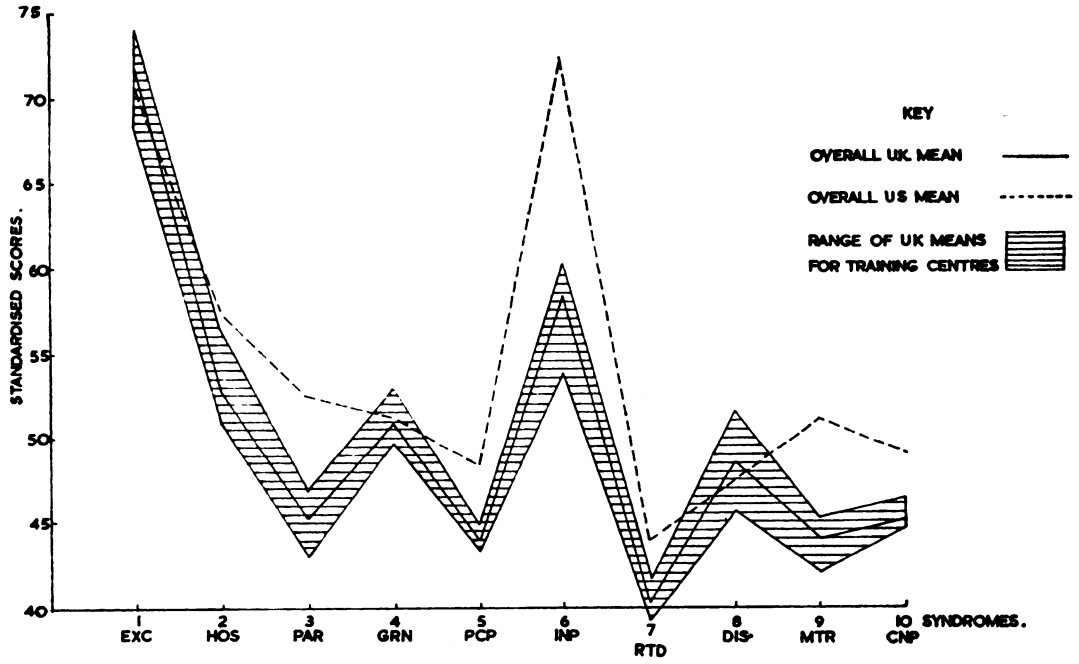

FIG. 3.—Inpatient multidimensional psychiatric scale. Mean symptom profiles of patient Mr. F for raters from U.K. centres of training.

https://doi.org/10.1192/bjp.118.547.629 Published online by Cambridge University Press

(giving ten syndrome scores for each of the tapes), the mean score for Maudsley-trained raters was lower than the mean score for all other British Isles raters, in 26 of the total of 30 syndrome scores. This difference reached a significant level in 13 of these 26 instances (seven being at a significance level of less than 1 per cent).

Dividing the raters into two groups by age (over 40 years and under 40 years) showed that there was a clear tendency for the older raters to rate higher than the younger. This was seen in its most striking form by a further comparison of those under 35 years with those over 49 years. Taking all three tapes together, the raters over 49 years old rated 22 of the possible 30 syndromes higher than the raters under 35 years, and this difference reached a significant level in 9 of the 22 instances. Six of these significant differences occurred in the tape of Mrs. A, the woman who described symptoms suggesting both affective and schizophrenic illness, and who obviously posed a difficult diagnostic problem.

Further examination of the ages of the various groups of raters shows that age and place of training are not independent. For instance, of the Maudsley-trained raters, 25 per cent are under 35 years but only 7 per cent over 50 years, as against 7 per cent and 17 per cent respectively for the raters trained in mental hospitals (the difference in syndrome score levels is greatest between these two groups).

This makes it necessary to examine the ratings for the effects of age and place of first four years of training separately. Table V sets out some of the necessary comparisons in summary form for the tape of Mrs. A. Similar

TABLE V

Differences in group mean syndrome scores of raters on the I.M.P.S. for age, and place of first four years of psychiatry. Videotape of Mrs. A.

(*Where the group mean syndrome score of the second named in each column is greater than the first, the significance of the difference by student's t-test, is shown by* (+) *for not significant,* + *for* $p < 0{\cdot}05$, ++ *for* $p < 0{\cdot}01$, +++ *for* $p < {\cdot}001$. *The converse is shown by a negative sign.*)

			All ages Maudsley (n = 44) vs. Mental Hospital (n = 82)	All under 35 years Maudsley (n = 11) vs. Other (n = 28)	All aged 35–49 years Maudsley (n = 30) vs. Other (n = 110)	Non-Maudsley under 35 years (n=28) vs. over 50 years (n=29)
IMPS	SYNDROMES	1	(+)	(+)	(+)	(+)
		2	+	(+)	(+)	(+)
		3	++	(+)	+	+
		4	(+)	(+)	(+)	(+)
		5	+	(−)	+	++
		6	+++	+++	(+)	(+)
		7	+++	(+)	++	(+)
		8	(+)	(+)	(−)	(−)
		9	(+)	+++	(+)	(+)
		10	(+)	+++	(−)	−−
			1	2	3	4

https://doi.org/10.1192/bjp.118.547.629 Published online by Cambridge University Press

but less marked differences are found for the other two tapes. Column 1 shows the difference between Maudsley and mental hospital trained raters in the most obvious form, since the effects of both age and training are present together. Column 2 compares Maudsley-trained raters under 35 years with all other raters in the same age group, and column 3 compares the middle-aged Maudsley (35–49 years) with the middle-aged non-Maudsley raters. It is evident that even when age is allowed for the Maudsley-trained raters still rate at a lower level than the others. This effect of place of training is greatest between psychiatrists trained at the Maudsley and those trained at mental hospitals, with those trained at other University Departments and undergraduate medical schools being intermediate. Column 4 shows that age alone has an influence amongst the non-Maudsley trained raters, although the differences here are not so obvious as in the other column. Only three Maudsley-trained raters were over 50 years, so it was not worthwhile comparing them with the younger Maudsley-trained raters. Insufficient raters were available from each centre of training to allow a full examination of the relative importance of these effects upon rating behaviour, but these examples serve to establish that at least two important influences, age and place of training, can be identified.

When considering either Mrs. A or all three tapes together, no particular syndrome is repeatedly affected, and the significant differences keep appearing for different syndromes. These are best regarded as very general rating tendencies which are likely to affect the rating of any type of symptom, given the conditions of these experiments.

Finally, examination of the type of work against the IMPS syndrome scores showed that psychiatrists whose main experience was in psychotherapy had a marked tendency to rate higher than the rest. Taking the syndrome scores for all three tapes, they rated higher in 24 of the 30 possible instances, and the difference reached a significant level in 5 of these. These two groups showed no significant differences in age.

(e) *Diagnoses and IMPS syndrome scores*

As might be expected, raters who make a diagnosis of schizophrenia also make higher ratings on certain syndromes than those who make other diagnoses. For Mrs. A, this is the case at a very significant level ($p < \cdot 01$ or less) on syndromes 3, paranoid projection; 5, perceptual distortion (which includes auditory hallucinations); and 10, conceptual disorganization (which includes items relevant to schizophrenic thought disorder).

In view of the previously noted tendency of Maudsley-trained raters to rate low, the association in American psychiatrists of high rating with a diagnosis of schizophrenia suggests the possibility that Maudsley raters would tend to avoid a diagnosis of schizophrenia. However, the proportions of diagnoses made according to the centres of training listed in Tables II–IV show that this is not the case. The Maudsley-trained raters had no special tendency to avoid a diagnosis of schizophrenia. This suggests that there is a specific low-rating tendency associated with a Maudsley training but not associated with any particular diagnostic habits, at least for these three tapes.

Discussion

These studies were undertaken because surprisingly little is known about the rating behaviour and diagnostic habits of psychiatrists. The need for more information is made particularly urgent by two developments in psychiatry in the last decade or so. First, effective psychotropic drugs are now available, but unless the patients upon whom drugs are tested are described and diagnosed by reliable and repeatable methods the results of therapy cannot be communicated to other workers with any confidence. Secondly, the fundamental importance of epidemiological studies is at last being recognized, and these also depend upon reliable and repeatable procedures for symptom rating and diagnosis.

In spite of the present state of comparative ignorance, there is no need for a particularly pessimistic outlook when considering the reliability of diagnostic procedures in psychiatry. It is easy to obtain a fairly gloomy overall picture from studies and reviews such as those

by Beck *et al.* (1962), Kreitman (1961) and Zubin (1967), but in the studies reviewed by these authors the diagnosticians were taking few precautions to avoid major sources of variation. More recent work (Wing *et al.*, 1967; Kendell *et al.*, 1968; Cooper *et al.*, 1969, 1970; Goldberg *et al.*, 1970) shows that when special steps are taken to avoid the most obvious sources of inter-observer variation, for instance by using standardized interview schedules and agreed glossaries of diagnostic terms, satisfactory levels of reliability and repeatability of the various stages of the diagnostic process can be achieved.

The IMPS schedule and associated procedures used in these studies represent a fairly crude level of standardization in that the raters were untrained and the interview was not geared in detail to the contents of the IMPS schedule. Nevertheless, the avoidance of technical psychiatric terms achieved by definition of the items in everyday language removes some of the major sources of variation between raters. We believe that the IMPS represents a satisfactory practical compromise which allows studies of rating behaviour to be done on a comparatively large scale without making impracticable demands on the raters' time.

The generalizations that can be made from this study are limited by both the size and nature of the audiences and the number of video-tapes. Since each patient had very different symptoms, it was not always possible to confirm a trend in the rating of a particular symptom across all three tapes or to study the effects of the presence of one group of symptoms on the perception of another. An additional problem is instanced by the finding that occasionally a symptom rated highly on two of the tapes was rated lower than average on the third by the same group of psychiatrists. Such unexplained inconsistencies between ratings on different tapes point to the dangers of assuming that the results from a single video-tape have necessarily a simple explanation or that they can be widely generalized. It must also be remembered that agreement between raters may be deceptive and due merely to patients showing either very slight or very marked degrees of a symptom. In spite of these cautions, we consider the findings here are important, since they are clear-cut and consistent and a large number of psychiatrists took part from all over the British Isles.

In this paper, discussion is inevitably centred around the differences between the psychiatrists, but this must not be allowed to obscure the overall conclusion that there is quite good diagnostic agreement between the various groups; in comparison to the differences between American and British raters, these differences within the British Isles are quite small. Of the British centres of training, Glasgow seems to be the most atypical. This confirms some of the findings of Sandifer *et al.* (1968) who, too, noted that raters in the Glasgow area differed in some of their diagnostic habits from a group of London psychiatrists. It cannot be said that any particular school of psychiatry is right or wrong in the diagnostic habits it inculcates, but it is clear that influential schools of teaching do exist and that they have lasting effects upon their trainees.

The differences that emerge between the various groupings of the psychiatrists on the IMPS syndrome mean scores are of particular interest, and none of them had been forecast. In view of the non-technical language of the IMPS ratings, it is perhaps surprising that the most consistent and clear-cut differences between psychiatrists occurred on IMPS syndrome scores rather than on diagnoses. The general level of rating is likely to be affected by the rater's attitude towards illness and health and what is normal and abnormal, and these findings show clearly that differences in such attitudes are associated with differences in psychiatric (postgraduate) training as well as with the more personal attribute of age. The low rating tendency of psychiatrists trained at the Maudsley Hospital and Institute of Psychiatry is general and fairly consistent, affecting the appreciation of all types of symptoms. No obvious explanation for this finding comes to mind: it is certainly not due to exposure to a large number of severely ill patients during the training period, which might have been put forward as a possible explanation had the psychiatrists trained in mental hospitals emerged at the bottom of the rating hierarchy. It could

be that this high rating threshold is produced by the academic orientation and the comparatively generous proportion of time spent on detailed clinical teaching that characterizes the Maudsley training programme, in that the trainees are influenced to become more careful and more conservative in their assessments.

The mere finding of these rating trends, however, is of great interest. It is important to know that they exist and to search for others, since the rating of symptoms is a necessary procedure in many types of research, as well as giving some indication of the judgements and perceptions concerned in the diagnostic process. It is also necessary to learn how to manipulate rating behaviour by varying both the nature of the ratings and the conditions under which the ratings are made.

Clinical teaching and practice will rest upon a more secure foundation when both pupils and teachers alike have some understanding of the influences that affect the perceptions and judgements underlying their diagnostic decisions.

Summary

Some 200 psychiatrists at seven different centres in the British Isles made diagnoses and ratings of symptoms and behaviour (using Lorr's IMPS) on three video-tapes of psychiatric interviews. Compared to large diagnostic differences demonstrated in other studies on the same video-tapes between British and American psychiatrists, there was quite good diagnostic agreement between the psychiatrists when grouped by their place of training. Only Glasgow stood out, in that psychiatrists who trained there had a significant tendency to make a diagnosis of affective illness in one of the tapes where the choice of diagnosis was between affective illness and schizophrenia.

On the IMPS ratings of symptoms and behaviour, Maudsley-trained psychiatrists had a significant general tendency to rate at a lower level than the rest. In addition, older psychiatrists rated significantly higher than younger psychiatrists, irrespective of their place of training, as did those with psychotherapeutic training and experience.

The comparative dearth of knowledge about the rating and diagnostic behaviour of psychiatrists is emphasized, and a plea is made for more studies of the effect of different methods and conditions upon these fundamental processes.

Acknowledgements

Special thanks are due to the many psychiatrists who made this study possible by attending the rating sessions, and to Officers and staff of the Royal Medico-Psychological Association who helped us in contacting them. We are also grateful to the Professors of Psychiatry at the Universities of Birmingham, Manchester, Glasgow, Edinburgh, Belfast and Dublin, and their staff, who provided facilities for the viewing sessions and gave generously of their hospitality.

This work was supported by Grant MH 09191 from the National Institute of Mental Health, U.S.A.

References

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. E., and Erbaugh, J. K. (1962) 'Reliability of psychiatric diagnoses: 2. A study of consistency of clinical judgements and ratings.' *Amer. J. Psychiat.*, **119**, 351–7.

Cooper, J. E., Kendell, R. E., Gurland, B. J., Sartorius, N., and Farkas, T. (1969). 'Cross-national study of diagnosis of the mental disorders: some results from the first comparative investigation.' *Amer. J. Psychiat.*, **125**, 10, Supp. 21–9.

—— (1970). 'The use of a procedure for standardizing psychiatric diagnosis.' In *Psychiatric Epidemiology*, eds. E. H. Hare and J. K. Wing, Oxford University Press.

Goldberg, D. P., Cooper, B., Eastwood, M. R., Kedward, H. B., and Shepherd, M. (1970). 'A standardized psychiatric interview for use in community surveys.' *Brit. J. prev. social Med.*, **24**, 18–23.

Greenhouse, S. W., and Geisser, S. (1959). 'On methods in the analysis of profile data.' *Psychometrica*, **24**, 95–112.

Gurland, B. J., Fleiss, J. L., Cooper, J. E., Kendell, R. E., and Simon, R. (1969). 'Cross-national study of diagnosis of the mental disorders: Some comparisons of diagnostic criteria from the first investigation.' *Amer. J. Psychiat.*, **125**, 10 Supp. 30–9.

——, ——, ——, Sharpe, L., Kendell, R. E., and Roberts, P. (1970). 'Cross-national study of diagnosis of the mental disorders: Hospital diagnosis and hospital patients in New York and London.' *Comprehensive Psychiat.*, **11**, 18–25.

H.M.S.O. (1968). *A Glossary of Mental Disorders.* Studies on medical and population subjects No. 22.

Katz, M. M., Cole, J. C., Lowery, H. A. (1969). 'Studies of the diagnostic process: The influence of symptom perception, past experience, and ethnic background on diagnostic decisions.' *Amer. J. Psychiat.*, **125**, 937–47.

KENDELL, R. E., EVERETT, B., COOPER, J. E., SARTORIUS, N., and DAVID, M. E. (1968). 'The reliability of the "Present State" Examination.' *Soc. Psychiat.*, **3**, 123–9.

——, SHARPE, L. COOPER, J. E., GURLAND, B. J., GOURLAY, A. J., and COPELAND, J. R. M. (1971). 'The diagnostic criteria of American and British Psychiatrists'. *Arch. gen. Psychiat.* (in press).

KREITMAN, N. (1961). 'The reliability of psychiatric diagnosis.' *J. ment. Sci.*, **107**, 876–86.

LORR, M., MCNAIR, D. M., KLETT, C. J., and LASKY, J. J. (1962). 'Evidence of ten psychiatric syndromes.' *J. consult. Psychol.*, **26**, 185–9.

SANDIFER, M. G., HORDERN, A., TIMBURY, G. C., and GREEN, L. (1968). 'Psychiatric diagnosis: A comparative study in North Carolina, London and Glasgow.' *Brit. J. Psychiat.*, **114**, 1–9.

SHARPE, L., GURLAND, B. J., FISHER, B., FLEISS, J. L. (1969). 'The accuracy of transatlantic communication in psychiatry.' Amer. Psychiat. Assoc. Meeting, Miami.

WING, J. K., BIRLEY, J. L. T., COOPER, J. E., GRAHAM, P., and ISAACS, A. D. (1967). 'Reliability of a procedure for measuring and classifying "Present Psychiatric State".' *Brit. J. Psychiat.*, **113**, 499–515.

ZUBIN, J. (1967). 'The classification of the behaviour disorders.' *Amer. Rev. Psychol.*, **18**, 373–401.

A synopsis of this paper was published in the *Journal* for October 1970.

J. R. M. Copeland, MB., M.R.C.P., D.P.M., *Senior Lecturer*,
John E. Cooper, B.M., M.R.C.P., D.P.M., *Consultant Psychiatrist*,
R. E. Kendell, M.D., M.R.C.P., D.P.M., *Reader*,
A. J. Gourlay, M.A., *Research Worker*,
U.S.–U.K. Diagnostic Project, Institute of Psychiatry, De Crespigny Park, London, S.E.5

(*Received 19 June 1970*)