

Expressed Emotion and Schizophrenia in North India

I. Cross-Cultural Transfer of Ratings of Relatives' Expressed Emotion

N. N. WIG, D. K. MENON, H. BEDI, A. GHOSH, L. KUIPERS, J. LEFF, A. KORTEN, R. DAY, N. SARTORIUS,
G. ERNBERG and A. JABLENSKY

A bilingual rater was trained in English in a technique of assessing relatives emotional attitudes to patients, and was then required to rate material in Hindi without any further experience. This strategy revealed that the rating of critical comments, hostility and positive remarks could be transferred from English to Hindi without distortion. There were problems with the remaining two scales, over-involvement and warmth, but these were due to technical issues connected with rating and not to cross-cultural distortion.

Measurement of expressed emotion

In order to study the emotional aspects of relationships between psychiatric patients and their relatives, a semi-structured interview, the Camberwell Family Interview (CFI), was developed in the 1960s (Brown & Rutter, 1966; Rutter & Brown, 1966). Originally this was administered to relatives at the time the patient was admitted to hospital.

The interview is tape-recorded, then rated from the recording. Ratings are made on five principal scales: *criticism*, *hostility*, *over-involvement*, *warmth* and *positive remarks*. *Criticism* is measured by the total number of critical comments made in the course of the interview. A critical comment is carefully defined in a manual used in conjunction with the CFI. *Hostility* is defined as either a generalisation of criticism (1), rejection of the patient as a person (2), or a combination of generalisation and rejection (3). *Over-involvement* can be detected and rated either from the respondent's behaviour during the interview (as with the other scales), or from reported behaviour outside the interview. It comprises several different elements, any one of which can give rise to a high rating. They are: exaggerated emotional responses, unusually self-sacrificing and devoted behaviour, over-protection that is inappropriate considering the patients' age, and an inability to maintain a boundary between the subject's existence and the patient's. *Over-involvement* is rated on a global scale extending from 0 to 5. *Warmth* refers not to a general quality in the respondent's personality, but to an expressed emotion specifically about the patient. It too is rated on a global scale which runs from 0 to 5. A *positive remark* is a statement which expresses praise, approval or appreciation of the behaviour or

personality of the patient. As with criticism, the number of positive remarks made throughout the interview is counted.

On all five scales, raters are required to pay attention not only to the content of relatives comments but also to vocal aspects of their speech, including rate, volume and tone. Despite the niceties of judgement demanded by this latter task, high levels of inter-rater reliability (about 0.8) are normally achieved after attendance at a training course of three to four weeks.

Implications of high EE

The value of this assessment has been established in a series of studies, reviewed by Kuipers (1979). Relatives were characterised as high on expressed emotion (EE) if they scored above certain levels for criticism or over-involvement, or if they expressed any hostility towards the patient. It was found that schizophrenic patients whose illness remitted and who were discharged from hospital to live with high-EE relatives were much more likely to relapse during a nine-month follow-up than those who returned to low-EE relatives (Brown *et al*, 1972; Vaughn & Leff, 1976a). These findings, originally derived from studies of families in London, have been replicated in a study in California (Vaughn *et al*, 1984); and evidence for a causal association between high EE and schizophrenic relapse has been provided by a controlled trial of intervention in high-EE families (Leff *et al*, 1982). In one of the earlier studies the degree of criticism expressed by relatives was also found to predict the outcome of neurotic depression (Vaughn & Leff, 1976a),

although the threshold for high and low criticism was much lower than that for schizophrenia. The finding has recently been replicated by Hooley *et al* (1986).

The importance of relatives' EE in psychiatric disorders is now beyond doubt in English-speaking cultures on both sides of the Atlantic. Considerable interest in this measure has also been expressed in other European cultures, and research workers from the Federal Republic of Germany, Holland, Austria and Belgium have been trained in this technique.

The WHO project on outcome of schizophrenia

In 1976 the World Health Organization initiated a multi-centre collaborative research programme on the outcome of schizophrenia, in which the EE measure has been incorporated. The project, on *Determinants of Outcome of Severe Mental Disorder*, is a prospective epidemiological, clinical and social study of schizophrenia and related conditions in different cultures, coordinated by the World Health Organization (Geneva). The study is financed jointly by WHO, NIMH, and twelve field research centres (FRCs) in Aarhus (Denmark), Agra and Chandigarh (India), Cali (Colombia), Dublin (Ireland), Honolulu and Rochester (USA), Ibadan (Nigeria), Moscow (USSR), Nagasaki (Japan), Nottingham (UK) and Prague (Czechoslovakia). All of the FRCs are participating in a 'core' epidemiological and clinical study and sub-groups of centres are carrying out a number of special sub-studies designed to test specific hypotheses about the course of schizophrenia, its social sequelae, and its associations with other diseases.

The 'core' study has set out to identify, over a 24-month period, all the psychotic patients (aged 15–54) making a first-lifetime contact with any 'helping agency' located in or near each of the twelve geographically defined catchment areas in the project. The clinical and diagnostic assessment of these individuals should permit an estimate of the incidence of schizophrenia and related syndromes. The one-year and two-year follow-up assessments of the patients should help to evaluate previous findings of differences in the prognosis of schizophrenia in the developing and developed countries (WHO, 1979).

In addition to the 'core' study, sub-groups of field research centres are undertaking certain special investigations. The one with which this paper is concerned focuses on the effects of patterns of emotional expression and communication in the family on the short-term course of schizophrenic illnesses.

The WHO study of EE and outcome

The centres chosen for a study of EE and outcome of schizophrenia were Aarhus in Denmark and Chandigarh in North India. Danish culture is similar to that in England, both being industrialised Western countries with a well-developed welfare system, whereas India provides a marked contrast. Furthermore, the two-year follow-up of the *International Pilot Study of Schizophrenia* (WHO, 1979) revealed that schizophrenic patients from certain developing countries, including India, had a better prognosis than those from centres in Europe and North America. This raised the possibility that more tolerant family attitudes to illness and handicap might explain or contribute to the better outcome in developing countries. A study of EE in Chandigarh could test this hypothesis.

In training candidates from other European countries to rate EE, we had assumed that the languages and cultures were similar enough for the technique to be transferable from English; but we considered that the cultural and linguistic differences between London and Chandigarh were too great for us to be comfortable with this assumption. The main language spoken in Chandigarh is Hindi, the tonal qualities of which are quite distinct from those of English. Since tone is a crucial element in rating EE, it was obviously essential to devise some method of assessing the transfer of the technique from English to Hindi.

Method

At first sight it may seem that routine tests of inter-rater reliability are sufficient to determine whether the assessment of EE can be transferred from one culture to another. However, a moment's consideration shows that this is not so. Suppose that two bilingual researchers are trained to an acceptable level of reliability in rating EE in English. They then return home and start rating in their native language. They check the reliability of their own ratings in their own language and find it is acceptably high. Unfortunately this does not prove the point. In the process of adapting the interviewing and rating techniques to their own cultural setting, it is inevitable that they will discuss between themselves the difficulties encountered. This could readily lead to a steady drift away from the techniques as taught. Because of their continued dialogue, the two raters could attain a high level of reliability in their own language whenever tested. However, this would by no means guarantee that what they were doing in their own language was the same as they had been taught to do in English. One way to check on this is to train another bilingual subject to rate EE reliably in English and then require him, before he has any experience of rating EE in the second language, to rate tapes of interviews made by the other raters in their native language. Since the 'naive subject', as we shall call him, has had no

opportunity to discuss the technique with the other raters, he has no option but to transfer the rating rules that he has learned in English to the second language as best he may. If he then reaches a high level of reliability with the other raters, it demonstrates that what they are doing in their native language must be very close to what was taught in English.

This was the approach we took to test the transfer to the interviewing and rating techniques for EE from English to Hindi. The local researchers (DKM and HB) were trained in English initially, one in London and one in Chandigarh, and achieved acceptable reliability levels on all scales compared with the London centre. The naive bilingual subject chosen was AG, who had no previous acquaintance with the techniques in either language. He was trained in English in the London centre to a satisfactory level of inter-rater reliability on the various scales. He was then required to rate tapes of interviews conducted in the Chandigarh centre in Hindi by the local researchers. It was originally planned that the whole exercise would be carried out on taped interviews sent from Chandigarh. Unfortunately some of the tapes were lost in transit, only six arriving safely in London. These were rated by AG as planned. As it happened, AG decided to visit Chandigarh at this time, and advantage was taken of this opportunity to sit in on four live interviews conducted by the local researchers and to rate them. AG fully understood the necessity to refrain from discussing interviewing or rating techniques with the local researchers until the exercise was completed, and he strictly observed this principle.

Results

Immediately following the training course, AG rated a standard set of tapes used to test reliability. He was found to have reached high levels of reliability on criticism, hostility and over-involvement, but not on warmth and positive remarks. Further attention was given to training AG and in early 1982, just before he visited Chandigarh, AG was asked to rate ten additional English tapes as a final check on his reliability. The product moment correlation coefficients between his ratings and those of the London centre researchers were as follows: critical comments 0.86, hostility 0.90, over-involvement 0.91, warmth 0.20, positive remarks 0.40. As before, his ratings of criticism, hostility and over-involvement were completely satisfactory, whereas those of warmth and positive remarks did not reach even a minimum standard of acceptability. This was not such a drawback as it might seem, since earlier studies had suggested that only the first three scales were associated with the outcome of schizophrenia, and only these three had been used to categorise relatives as high- or low-EE. However, we considered it important to measure these elements of EE in the Chandigarh relatives, since it was feasible that their relationship with outcome might differ considerably from that found in the Anglo-American studies.

The problem AG experienced in becoming reliable on the warmth and positive remarks ratings was unexpected, as trainees are usually able to master all the EE scales by the end of the course. Moreover this problem failed to yield to a further period of training, but because of AG's imminent

departure for Chandigarh, we were not able to try other approaches.

The raw scores on the Hindi interviews rated by AG and the two Chandigarh raters (DKM and HB) are displayed in Table I. The data from the six taped interviews which were sent to England have been combined with those from the four interviews at which AG was present. An inspection of this table reveals that the three raters differ little in ratings of critical comments, hostility and positive remarks; whereas for over-involvement and warmth differences are more marked.

For critical comments the intraclass correlation coefficient (ICC) for the three raters is 0.87, a satisfactory value. There are some differences between the raters but these are not consistent and are small compared to differences between the subjects. A positive score for hostility was recorded on only four of the thirty assessments. However, there was complete agreement between the three raters on the presence or absence of hostility in nine of the ten cases. There was considerably more disagreement on emotional over-involvement. The intraclass correlation coefficient (0.21) is not statistically significant. We conclude that the three raters were rating differently, and a visual inspection of the results shows that AG consistently rated higher than either DKM or HB, who differed little from each other. The ICC for warmth is only 0.49, which is barely significant. Once again, AG rated differently from DKM and HB, giving a lower rating on all but one subject. The agreement between the three raters is satisfactory for positive remarks, the ICC being 0.86.

The current criteria for assigning relatives to a high-EE category are: (1) six or more critical comments, or (2) any positive rating on hostility, or (3) a score of three or more on over-involvement. When these criteria are applied to the material in Table I, there is complete agreement between the three raters on the assignment of two relatives to a high-EE category and seven relatives to a low-EE category. In one case (number 7) there is a disagreement, with the two Chandigarh raters scoring the subject low on critical comments while the naive rater scored above the cut-off point for a high-EE rating.

Discussion

It is evident from these results that the rating of critical comments can be transferred satisfactorily from English to Hindi. This is reassuring, since critical comments contribute most to the assessment of EE. In Vaughn & Leff's (1976a) study, 18 out of 21 families (86%) were assigned to a high-EE group on the basis of six or more critical comments.

We cannot be so certain about the transferability of ratings of hostility since so few subjects in the Chandigarh reliability exercise were given a positive rating on this scale. However, the findings suggest that this assessment is also likely to be transferable across the linguistic frontiers without significant distortion.

The position regarding over-involvement is less

TABLE I
Raw scores of individual raters

Subject	Critical comments			Hostility			Over-involvement			Warmth			Positive remarks		
	DKM	HB	AG	DKM	HB	AG	DKM	HB	AG	DKM	HB	AG	DKM	HB	AG
1	2	2	2	0	0	0	0	0	1	3	3	2	2	1	2
2	12	14	14	1	2	2	0	0	0	1	1	0	0	0	0
3	0	0	1	0	0	0	1	1	1	3	3	3	4	4	4
4	2	2	2	0	0	0	0	0	1	3	3	2	1	1	1
5	4	4	4	0	0	0	0	0	0	3	3	2	0	0	0
6	1	2	2	0	0	0	0	0	0	2	2	1	1	1	1
7	3	1	8	0	0	0	0	0	0	0	1	1	0	1	1
8	0	1	4	0	0	0	0	2	2	3	2	1	2	2	2
9	1	0	1	0	0	0	0	0	2	3	4	1	1	1	2
10	12	15	9	0	1	0	0	0	1	2	2	1	0	1	2
Mean	3.7	4.1	4.7	0.1	0.3	0.2	0.1	0.3	0.8	2.3	2.4	1.4	1.1	1.2	1.5
Intraclass correlation coefficient	0.87			0.79			0.21			0.49			0.86		

satisfactory. AG noted some degree of over-involvement in a higher proportion of subjects than did DKM, and HB was much closer to her Chandigarh colleague in this respect than to AG. Since AG showed no tendency to over-rate on this scale in English, we can conclude that the Chandigarh researchers were likely to be under-rating in Hindi. However, this is not a consequence of cultural differences, as far as we can judge from AG's observation of the live interviews in Chandigarh and discussions with the researchers there.

He reported that judgements of over-involvement were almost exclusively based on overt expression *during the interview* of over-concern or excessive anxiety about the patient's condition. Relatively little attention was paid to reports of behaviour by the subject that reflected over-involved attitudes. This clearly represents a shift of emphasis subsequent to training: during training raters are instructed to give equal weight to the expression of emotional over-involvement and reports of behaviour that suggest it. This drift away from the original convention apparently affected both raters and is likely to have led to under-estimates of over-involvement in the Chandigarh sample of relatives.

The possibility of a drift away from the rating conventions inculcated during a training period has to be guarded against with any rating scale. A number of precautions can be taken. It is very helpful to have a detailed manual of rating instructions available, which trainees can take away with them: this is the case with EE assessment. If there are two trainees in one centre, they should be instructed to carry out inter-rater reliability checks at regular intervals. Finally, trainees should send a proportion of their taped interviews, say one in ten, to the training centre for monitoring by the trainers. This creates extra work for the training centre but is essential for quality control of rating techniques. Of course in the case of Chandigarh it was not possible for the trainers to assess the reliability of ratings of interviews conducted in Hindi. Only the technique of using a 'naive' bilingual rater can reveal the kind of drift that we detected in respect of the over-involvement scale.

Our assessment of the reliability of the transfer of warmth ratings from English to Hindi was clouded by the difficulty AG experienced in mastering this scale in English. He consistently tended to under-rate warmth when assessing English tapes, and the same phenomenon is evident in his Hindi ratings, since both

Chandigarh raters scored higher than he did for the subjects in this study. AG's relatively low reliability with the Chandigarh raters on the warmth scale is almost certainly due to his difficulty with these ratings in both languages. Whether there is an additional problem in transferring this assessment from English to Hindi cannot be determined from these data.

The results for the 'positive remarks' scale are peculiar, since AG attained a reliability on this scale of no more than 0.40 with the English raters, but reached highly satisfactory levels of agreement with the Chandigarh workers in terms of the intra-class correlation coefficient. How is it possible for him to have achieved so close a unanimity with the Chandigarh raters when he appeared to have difficulties with this scale in English? We can only speculate that he grasped the principle behind the rating of positive remarks but was unable to operationalise it satisfactorily in English. However, when he applied it in Hindi without any prior experience of using it in this language he employed it in a very similar way to the other Hindi-speaking raters.

The findings of this study leave us reasonably satisfied that the EE components of critical comments, hostility and positive remarks can be transferred intact from English to Hindi. In the case of over-involvement it was discovered that a drift had occurred away from the rating conventions established during the training. A drift of this kind could just as easily have occurred with researchers rating in

English, and does not indicate a specific problem in transferring the rating procedure across languages. However, the presence of this rating bias prevented us from determining whether an additional cross-cultural distortion existed. The difficulties the bilingual rater experienced in achieving reliability on the warmth scale were equally evident in English and Hindi: they therefore represent an idiosyncratic problem and not a cross-cultural one.

We conclude that the strategy of using a 'naive' bilingual rater is an effective method of evaluating the cross-cultural transfer of a rating technique. However it is clear that this strategy is viable only if the naive rater is able to become highly reliable on all the scales to be compared across cultures. This exercise was undertaken to assess the validity of comparing the effects of relatives' EE on the course of schizophrenia in Chandigarh with previous work in English-speaking cultures. We have complete confidence in the ratings of critical comments in Hindi and reasonable confidence in the hostility scores, but some doubts about the ratings of over-involvement. These are the three components which have contributed to the assessment of high-EE relatives in the past. Insofar as critical comments have made the major contribution to the detection of high EE, we consider that a direct comparison between the findings of the Chandigarh study of EE and the work in Britain and North America is meaningful.

II. Distribution of Expressed Emotion Components among Relatives of Schizophrenic Patients in Aarhus and Chandigarh

N. N. WIG, D. K. MENON, H. BEDI, J. LEFF, L. KUIPERS, A. GHOSH, R. DAY, A. KORTEN, G. ERNBERG,
N. SARTORIUS and A. JABLENSKY

We measured the components of expressed emotion among two samples of relatives of first-contact patients from Aarhus (Denmark) and Chandigarh (India). The Danes were very similar in most respects to samples of British relatives, whereas the Indian relatives expressed significantly fewer critical comments, fewer positive remarks, and less over-involvement. Within the Chandigarh sample, city-dwellers were significantly more expressive than villagers of all EE components except over-involvement.

As stated in the first paper of this series of three, a study of the relationship between relatives' expressed emotion (EE) and outcome of schizophrenia was undertaken in Aarhus in Denmark and Chandigarh in North India. It was assumed that the techniques of

interviewing and rating were readily transferable from English to Danish, these being closely related languages. The same assumption could not be made about English and Hindi, the predominant language in Chandigarh.