The Reliability and Validity of Measures of Family Life and Relationships in Families Containing a Psychiatric Patient

MICHAEL RUTTER and GEORGE W. BROWN

Institute of Psychiatry, Maudsley Hospital, London

Summary. Several studies have highlighted the burden on the families of mentally ill patients and so emphasised the need to investigate the effects of illness on family life and relationships. A review of the literature showed that many of the published measures of the family lacked satisfactory reliability and validity. — A methodological study of 30 families in all of which there were children of school age or younger and in which one parent had newly attended or was being treated at a psychiatric facility is reported. For each family there was a 2--3 hour interview with the patient, a 3--4 hour interview with the spouse and a third interview of about an hour with the two of them together. At each interview there were two investigators who made independent ratings. There were different investigators for each of the three interviews so that there were 6 investigators and 6 sets of independent ratings for each family. — A basic distinction was made between concrete happenings or activities in the family, on the one hand and, on the other, feelings, emotions or attitudes concerning these events or the individuals participating in them. An interview designed to measure both types of variables is described. By the use of a particular kind of flexible and detailed questioning about recent events an attempt was made to free the measurement of events and activities from attitudes and to develop different approaches for each. — The use of tape recordings and group discussions, interviewers were trained to recognise and rate differences in tone of voice as shown in the speed, pitch and intensity of speech. The third main type of scale consisted of overall summary ratings of relationships. — Each type of scale is illustrated by a number of examples for which detailed results are given. These are 'warmth' and 'number of critical remarks' as examples of 'emotional' scales, 'participation of husband and wife in household tasks' and 'frequency of irritable acts' as examples of scales of events or happenings, and 'overall assessment of the marriage' as a summary scale. Details are also given on the scale measuring 'parent preoccupations' which does not readily fit into any of these 3 main categories. — High inter-rater reliability was demonstrated for the emotional scales. Respondent and rater biases were not found. By comparison of ratings made in different interviews, it was shown that similar feelings were expressed about the spouse in different situations, suggesting that the ratings were valid as well as reliable. A person's report of someone else's emotions, on the other hand, appeared to have a low level of validity in spite of good reliability. — Inter-rater reliability for the measures of activities and for the overall summary ratings was also high and, more important, it was shown that there was good agreement between the accounts of husbands and wives when interviewed separately. — A detailed examination was made of possible factors influencing husband-wife agreement. Slight biases associated with dissatisfaction and patient status were found but these were not sufficient to influence conclusions in most instances. Ways of reducing such bias are discussed.

Résumé. Plusieurs études ont mis en lumière la charge que représente pour les familles un malade mental; elles ont fait ressortir la nécessité d'enquêter sur l'effet de la maladie sur la vie familiale et les relations. Un examen de la littérature a montré que les renseignements fournis par les familles manquaient d'objectivité et de valeur. Une étude méthodologique a été entreprise en tenant compte des problèmes énumérés. Pour chaque famille il y avait des enfants d'âge scolaire ou plus jeunes, et dont un des parents avait été récemment admis en établissement psychiatrique, a été fait. — Chaque famille a fait l'objet de 2--3 heures d'inter- view avec le patient, de 3--4 heures d'entretien avec le conjoint, et d'une interview avec l'enfant, sur la base de rapports réunis. Lors de chaque entretien il y avait deux enquêteurs qui faisaient chacun un rapport séparé, en sorte qu'il y avait en tout pour chaque famille six enquêteurs et six différentes appréciations. — Une distinction a été faite entre événements concrets et activités de la famille, d'une part, et, d'autre part, sentiments, émotions et attitude concernant ces événements ou la participation individuelle. Par un questionnaire souple et détaillé sur les événements récents, on a tâché de dégager l'appréciation des événements et des activités de tous préjudiciable. Une approche différente a été employée pour apprécier les émotions. Par le moyen de magnétophones et de groupes de discussion, les enquêteurs ont été entraînés à reconnaître et mesurer les différences de ton de voix selon la vitesse, le diapason, l'intensité du langage. Le troisième échelon a consisté en apprécier les appréciations sur des relations et de manière globale, la validité de l'échelle de degrés est illustrée par des exemples pour lesquels des résultats détaillés sont donnés. Il y a « chaleur » et « nombre de remarques critiques » comme exemples du degré d'affection ; « participation du mari et de la femme aux tâches ménagères » et « fréquence des actes d'irritation » comme exemples d'événements ; et « conception générale du mariage » comme resumé. On donne également des détails sur le plan « parent que l'enfant préfère », appréciation qui ne trouve pas facilement sa place dans une des trois catégories. — Quant au plan affectif, la comparaison a montré une grande justesse des appréciations. On n'a pas noté de parti-pris chez les enquêteurs, ni chez les référents. La comparaison des appréciations faites en différents entretiens a montré que des sentiments identiques étaient exprimés au sujet du conjoint dans différentes situations, ce qui permet de penser que les appréciations des enquêteurs étaient objectives. Par le moyen de magnétophones et de groupes de discussion, interviewers were trained to recognise and rate differences in tone of voice as shown in the speed, pitch and intensity of speech. The third main type of scale consisted of overall summary ratings of relationships. — Each type of scale is illustrated by a number of examples for which detailed results are given. These are 'warmth' and 'number of critical remarks' as examples of 'emotional' scales, 'participation of husband and wife in household tasks' and 'frequency of irritable acts' as examples of scales of events or happenings, and 'overall assessment of the marriage' as a summary scale. Details are also given on the scale measuring 'parent preoccupations' which does not readily fit into any of these 3 main categories. — High inter-rater reliability was demonstrated for the emotional scales. Respondent and rater biases were not found. By comparison of ratings made in different interviews, it was shown that similar feelings were expressed about the spouse in different situations, suggesting that the ratings were valid as well as reliable. A person's report of someone else's emotions, on the other hand, appeared to have a low level of validity in spite of good reliability. — Inter-rater reliability for the measures of activities and for the overall summary ratings was also high and, more important, it was shown that there was good agreement between the accounts of husbands and wives when interviewed separately. — A detailed examination was made of possible factors influencing husband-wife agreement. Slight biases associated with dissatisfaction and patient status were found but these were not sufficient to influence conclusions in most instances. Ways of reducing such bias are discussed.

The importance to the psychiatrist of investigations of family life and relationships no longer needs arguing. Much of psychiatric theory is concerned with the role of abnormal family structure or deviant parental attitudes and behaviour in the genesis of psychiatric disorder. Stemming from observations that mental illness often affects several people in the same family at about the same time (Post and Wardle, 1962), psychiatrists have become increasingly concerned with the diagnosis and alleviation of family psychopathology in contrast to the previous more exclusive concern with the individual patient (e.g. Ackerman, 1958; Bell, 1962; Howells, 1963; Laing and Esterson, 1964). Brown (Brown et al., 1962) has shown that family attitudes and relationships may influence the course of schizophrenia, and recently too, evidence has accumulated which points to the adverse effects of long-standing illness in a parent on the health of the rest of the family (Rutter, 1966). Investigations by Clausen and Yarrow (Clausen and Yarrow, 1955; Clausen, 1959), Grad and Sainsbury (Sainsbury and Grad, 1962; Grad and Sainsbury, 1963), and Brown and Wing (Wing et al., 1964; Brown et al., 1966) have all emphasised the extent of the burden on the families of mentally ill patients.

Mental illness may have deleterious effects on the rest of the family through the direct social effects of certain symptoms, by alterations in the balance of family activities and structure and through adverse effects on interpersonal relationships (Rutter, 1966). If these effects are to be investigated effectively, it is essential to have reliable and valid tools for the measurement of different aspects of family life and relationships, and it is with this problem that the present paper is concerned.

Review of Literature

The volume of research on the family over the last 40 years has been considerable (Nimkoff, 1948; Nye and Bayer, 1963), but other writers have pointed to the methodological weaknesses of much of the work (Hoffman and Lippitt, 1960; Nye and Bayer, 1963; Yarrow, 1963). Consequently in this brief review of the literature we will comment only on a few investigations which have been concerned with methodological issues and in which the findings illustrate some of the major problems.

Issues involved in the assessment of the reliability and validity of measures of family life have been considered by a number of writers (Cuming et al., 1958; Dean and Whyte, 1958; Hoffman and Lippitt, 1960; Vidich and Bensman, 1954), others have compared the value of interview, observational, and questionnaire techniques (Ellis, 1947; McCord and McCord, 1961; Richardson et al., 1965; Smith, 1958; Yarrow, 1963) and the relative merits of retrospective and longitudinal approaches have also been discussed (Baldwin, 1962; Bell, 1960; Yarrow, 1963; Yarrow et al., 1964). A basic distinction must be made between concrete happenings or activities in the family on the one hand and on the other, feelings, emotions, or attitudes concerning these events or the individuals participating in them (Clark and Sommers, 1961). Dean and Whyte (1958), Garrett (1942) and Hoffman and Lippitt (1960) made a similar distinction between objective and subjective items, and Yaukey and his colleagues (1965) between facts and feelings. Approaches to the assessment of validity are necessarily quite different for these two types of measure.

With events and activities there is an objective reality to be measured. Regardless of people's feelings about it there are factual aspects to the number of contacts the family have had with kin or with friends, the amount the husband helps with the housework, the frequency of quarrelling etc. With such items the concern is to eliminate any distortions and biases involved in the reporting of the informant in order to get as close as possible to the objective reality. It is, of course, necessary that the report be reliably scored or rated but the main issue is the accuracy of the report and the major way this may be tested is the comparison of the informant's account with the accounts given by others (Dean and Whyte, 1958).

With feelings and attitudes the situation is quite different. Comparison of different accounts is unhelpful. The wife's account of her husband's attitudes in no way validates his own account. Here the problem is to get the informant to express his attitudes in a way that does not distort his inner feelings and then to get investigators to agree on their ratings. Similarly the main problem in the assessment of emotional states such as warmth or hostility is the inter-judge agreement on what is observed. In contrast to the situation with objective happenings, ambiguities and ambivalence are frequent with feelings. Someone may feel both warmth