

Brief measure of expressed emotion: internal consistency and stability over time

SEIJA SANDBERG¹, MICHAEL RUTTER², JOHANNA JÄRVI³

¹ Department of Psychiatry and Behavioural Sciences, Royal Free and University College Medical School, London, UK

² Institute of Psychiatry, Kings College London, UK

³ Department of Psychology, University of Jyväskylä, Finland

ABSTRACT *The study examined three methodological aspects of expressed emotion (EE) as assessed in the course of PACE (Psychosocial Assessment of Childhood Experiences) interviews with a parent. In a sample of 87 children, aged 6–13 years, enrolled in a prospective study examining the role of stress on the course of asthma, EE was assessed at three time points, 9 months apart. A high degree of agreement was found among the three concurrent measures of negative and positive EE (kappas from 0.74 to 0.97, and from 0.45 to 0.88, respectively; $p \leq 0.0001$ in all instances). The temporal stability of all measures was lower, although statistically significant in all but 2 instances (kappas from 0.19 to 0.59, and from 0.11 to 0.39, respectively). The temporal stability across measures, as well as across interviewers and over time, was broadly similar (kappas from 0.21 to 0.56 for negative EE, and from 0.09 to 0.38 for positive EE, with all but three of the 36 statistically significant). The findings provide support for the underlying assumptions of the PACE-EE and show the utility of measures based on just very brief periods of non-directive interviewing, making them practical in a wide range of studies with EE just one of a larger set of measures.*

Key words: expressed emotion, PACE, children, temporal stability, across-informant agreement

Introduction

The stimulus for the development of measures to assess the qualities of relationships within families arose from both studies of the families of patients with schizophrenia (Brown et al., 1962) in which the focus was on the emotions expressed by relatives towards the patients, and studies of families in which one parent had a mental disorder (Rutter, 1966) in which the focus was on the emotions expressed by the parent-patient towards their children. This led to the development of the Camberwell Family Interview (CFI) and to measures of warmth and of severity of criticism (Brown and Rutter, 1966; Rutter and Brown, 1966) as well as to counts of the number of positive and critical comments. The measures were innovative in four key respects:

- they focused on individual-specific expressed

emotions (individual with respect to both the person expressing and the person receiving) rather than general family qualities;

- they used emotions as observed in the manner an individual talked about another person, rather than answers to specific closed questions;
- they used both verbal and vocal elements in rating emotions (that is, both what was said and the tone of voice used); and
- they focused on emotions expressed about a person as an individual, rather than those concerned with feelings about that person's symptoms.

Despite dealing with subtle aspects of feelings, the measures were found to have a high degree of inter-rater reliability and to agree well with the emotions observed when the two people interacted. In other

words, the way in which one person talked about another person was found to be predictive of the way in which they actually interacted with that person.

Since those early studies nearly four decades ago, the role of expressed emotion (EE) has been investigated extensively in relation to psychiatric and physical conditions both in adults (Miklowitz et al., 1984; Hahlweg et al., 1989; Kavanagh, 1992; Nugter et al., 1997), and in children (see Hodes et al., 1999b; Liakopoulou et al., 2001; Wamboldt et al., 1995, 2000, for somatic illnesses; and see Schwartz et al., 1990; Hibbs et al., 1991, 1993; van Furth et al., 1993; Asarnow et al., 1993, 1994; Vostanis, Nicholls and Harrington, 1994; Hirshfeld et al., 1997; Peris and Baker, 2000 for psychiatric conditions). Whereas the majority of studies have examined the prognostic power of EE in the context of adult psychopathology, its utility with regard to childhood disorders has been of more recent interest.

Traditionally, EE ratings have been based on the lengthy semi-structured Camberwell Family Interview (CFI) (Brown and Rutter, 1966; Rutter and Brown, 1966; Brown et al., 1972; Vaughn and Leff, 1976) that focuses on the 3 months prior to hospital treatment in order to explore the features of criticism, hostility and intrusiveness. More recently, Hodes et al. (1999a,b) have adapted the CFI for use in studies concerning children; and Scott and Campbell (2001) have produced a shorter version. Other studies with children have employed the Five Minute Speech Sample (FMSS) developed by Magaña et al. (1986) as a less time-consuming method of gleaning similar information. The measure requires relatives to speak for 5 minutes without interruption, characterizing their relationship with, and attitude towards the patient. The underlying notion is that the pressure elicited by a time constraint is likely to evoke core emotions more readily, and that the stimulus should be non-directive in order to avoid focusing on psychopathology. The concurrent validity of the FMSS has been established (Magaña et al., 1986; Leeb et al., 1991; Malla et al., 1991; van Furth et al., 1993), and much of the child EE research has relied upon ratings obtained from the FMSS.

Criticism and emotional over-involvement (EOI) concern the most common elements of EE, especially in studies with adults. Of these two, criticism taps feelings of negativity or resentment directed towards

the family member concerned, and EOI provides an index of behaviours that may be overprotective or self-sacrificing in the extreme. Other dimensions included in the EE construct are warmth, positive comments and hostility. Studies carried out especially with younger children have, however, highlighted problems with the measure of EOI both in terms of its uncertain validity and rare occurrence (for example, Baker et al., 1998, quoted by Peris and Baker, 2000; Hodes et al., 1999a,b).

The attitudes addressed by EE measures have been shown to correlate with the degree of criticism or overbearing commentary made during actual family interaction (Miklowitz et al., 1984). Hodes et al. (1999a) also demonstrated a moderately high agreement between EE ratings on the CFI and those based on a structured family interview applied in clinical practice.

Reliability of EE measurement

The reliability of EE has been established with satisfactory test-retest and inter-rater agreements (Malla et al., 1991; Barnes-McGuire and Earls, 1994; McGuire and Earls, 1994; Hodes et al., 1999a, b; Tarrier, Sommerfield and Pilgrim, 1999). Hodes et al. (1999b), for example, in a study of children and adolescents suffering from epilepsy, reported inter-rater reliabilities (Spearman correlation coefficients) of 0.96 for positive remarks, 0.80 for warmth, 0.94 for critical comments and 0.90 for emotional over-involvement. The reliabilities reported by Scott and Campbell (2001) were similarly high. The figures in a study of 43 adult post-traumatic stress disorder (PTSD) sufferers by Tarrier et al. (1999) were very similar, with a 100% agreement on overall EE status, and correlations of 0.75 for critical comments, of 1.0 for hostility, of 0.97 for warmth, of 0.55 for positive remarks and of 0.95 for EOI.

Agreement between different components of the EE construct

The Tarrier et al. (1999) study also examined correlations between the overall EE measure and individual scales. The overall EE rating was found to be significantly ($p < 0.001$) correlated with critical comments ($r = 0.68$), hostility ($r = 0.66$), and EOI ($r = 0.64$), but not with warmth or positive comments. In addition, critical comments were highly correlated with hostility ($r = 0.77$, $p < 0.001$), moderately correlated with EOI

($r = 0.35$, $p < 0.05$), and showed a moderate negative correlation with warmth ($r = -0.38$, $p < 0.05$), but no significant correlation with positive comments. Hostility was negatively correlated with warmth ($r = -0.39$, $p < 0.05$), but not with EOI or positive comments. EOI, warmth and positive comments showed no significant associations with each other.

Stability of EE

The research examining the stability of EE in adults has produced somewhat inconsistent findings. For example, Patterson, Birchwood and Cochrane (2000) in a study of the key relatives of 50 early onset, first-episode psychotic patients examined the natural history of EE and found that EE status altered in over a quarter (28.2%), primarily from high to low EE. The ratings derived from the FMSS have, on the other hand, demonstrated significant stability over periods up to 5 years in other studies (Leeb et al., 1991; McCreddie et al., 1993; King, 1998).

Few studies have so far explored the stability of EE with regard to children. In an effort to examine the test-retest reliability of the FMSS, McGuire and Earls (1994) administered the measure to parents of 29 disadvantaged children. Expressed emotion ratings remained moderately stable over a 5-week period. It was therefore suggested that such findings support the notion of EE being a parental trait rather than a reflection of mood or circumstances (Schreiber, Breier and Pickar, 1995). Scott and Campbell (2001) reported kappa values of 0.36 to 0.42 for stability over a 5-month period. A study by Peris and Baker (2000) also demonstrated considerable stability of EE as measured by the FMSS. This was especially the case with high parent-expressed criticism, and its relationship to disruptive behaviour in a community sample of preschoolers. The stability of the EE rating over a 2-year period from pre-school to first grade (from age 4 to 6 years approximately) was statistically significant, although modest.

Differentiations between disturbed and non-disturbed children

There is fairly consistent evidence from studies carried out so far that children suffering from emotional and behavioural problems are subject to higher levels of parent-expressed negative emotion, especially criticism, compared with their non-

disturbed peers (Hibbs et al., 1991, 1993; Stubbe et al., 1993; Van Furth et al., 1993; Kershner, Cohen and Coyne, 1996; Hirshfeld et al., 1997; Baker, Heller and Henker, 2000; Scott and Campbell, 2001). In addition, there is some indication that particular aspects of parent EE differentiate among individual child psychiatric disorders (Schwartz et al., 1990; Stubbe et al., 1993; Asarnow et al., 1994; Vostanis et al., 1994).

Prognostic validity of EE in childhood disorders

Apart from concurrent validity, the EE construct also appears to show prognostic validity with children regarding the course of both psychiatric (Schwartz et al., 1990; Asarnow et al., 1993; Vostanis and Nicholls, 1995; Baker et al., 2000; Peris and Baker, 2000), and somatic conditions (Wamboldt et al., 1995, 2000; Hodes et al., 1999b).

PACE-EE

The 5-minute speech sample had the major advantage of brevity, making it readily applicable in a wide range of studies. However, it required the informant to talk for 5 minutes without prompting, and this involved two substantial disadvantages. First, in our experience, many people (perhaps especially the less well educated and, for different reasons, those whose first language was not English) found the requirement unpleasant, and tended to 'freeze'. Second, it did not fit well into the conversational style of an investigator-based interview and so interfered with rapport.

Accordingly, as part of PACE (Psychosocial Assessment of Childhood Experiences, Sandberg et al., 1993), we developed a modified version that involved the same principle of eliciting free description without focused probes, but which differed in allowing seven specified forms of non-directive encouragement to help the informant extend their description (see Appendix 1).

As with other methods of eliciting EE, the ratings of the expressed emotions have to be made by investigators trained to do so reliably, following the criteria as originally developed for the Camberwell Family Interview (Brown and Rutter, 1966; Rutter and Brown, 1966; Brown et al., 1972; Vaughn and Leff, 1976; Leff and Vaughn, 1985). This modified version, the PACE-EE, has been used in a variety of

Institute of Psychiatry (London) epidemiological studies involving families with children during the 1980s to 1990s (for example, Rutter and Quinton, 1984; Taylor et al., 1991).

Aims of the present study

The present study set out to examine three methodological aspects of EE as assessed in the course of PACE interviews with a parent. First, it was necessary to determine the degree to which the PACE-EE ratings based on the brief period (approximately 5 minutes) of non-directive prompts agreed with the same ratings made on the total interview (lasting some 1 to 2 hours). Second, the stability over time (between interviews 9 months apart) was examined using different interviewers for the two interviews (in order to ensure independence); and third, temporal stability was similarly examined with the additional consideration of using the brief assessment from one interview but the full interview for the other. All of these were examined in a non-psychiatric sample in order to assess the qualities of measured EE in a group of families with little psychopathology.

Methods

Subjects

The sample consisted of 87 children (female 46%, male 54%), aged 6–13 years, enrolled in a prospective study examining the role of stress on the course of chronic asthma using the PACE standardized interview, which includes the assessment of expressed emotion (see below). The only inclusion criteria were the child's age and asthma status. Of the total of 90 children included in the study, three were omitted from the present analyses due to incomplete EE data at any of the three time points. The sample has been fully described in Sandberg et al. (2000). The prime purpose of the study was to use a prospective design, with measurement at baseline and at 9-month and 18-month follow-ups, to examine the role of psychosocial factors on the course of asthma. As EE was assessed using the whole interview and brief method based on the modified 5-minute speech sample, the opportunity was taken to examine the agreement across the two methods. Similarly, the repeat measurement of EE provided

the opportunity to examine temporal stability both within and across measures.

Methods

The assessment of expressed emotion was carried out as part of the investigator-based PACE interview with the parent at baseline, and at two subsequent time points 9 months apart. Two separate interviewers were involved, thus allowing for 'cross-over' of interviewers with person A carrying out baseline assessment, person B the 9-month follow-up, and person A the 18-month follow-up, or in the reverse order.

The PACE is a standardized, investigator-based interview with established reliability and validity (Glen et al., 1993; Sandberg et al., 1993). The interview covers a wide range of psychosocial experiences, acute and chronic, relevant for the lives of children. The PACE has parallel versions, one for the use with parents and one with the use with children. The data used in the present analyses are based on the information obtained in the parent interviews.

Expressed emotion assessment in PACE comprises three main ratings: namely warmth, criticism and hostility, each being rated on a four-point scale extending from '0' to '3'. Hostility was omitted from the analyses reported here because of its very rare occurrence in the non-psychiatric sample used. The ratings of warmth and criticism were made in the same way in the brief period of non-directive prompts (termed PACE-EE for short) and during the whole interview (termed WI for short). The whole interview was also used to provide a total count of positive comments (No. Pos. Comm.) and of critical comments (No. Crit. Comm.). Unlike warmth and criticism, these are based solely on the content of what is said.

The training for the PACE involves both training in how to administer the interview correctly and training in the rating procedures. These procedures involve detailed guidelines on verbal and vocal elements to be used in ratings and guidelines on the criteria for different points on the scale.

The level of agreement in ratings over time and between measures was examined using Cohen's (1968) weighted kappa statistic, which was designed to assess agreement between categorical variables, correcting for chance agreement as expected from

the base rate distribution of the ratings in the two sets of ratings to be compared (see Kraemer, Periyakoil and Noda, 2002).

Results

Agreement in terms of kappas was calculated for negative expressed emotion as based on the above three measures separately – PACE-EE, Whole Interview-EE (WI-EE) and Number of Critical Comments (No. Crit. Comm.), with the results shown in Table 1.

There was a high degree of agreement among the three concurrent measures of negative expressed emotion, with kappas ranging from 0.74 to 0.97, all significant at $p \leq 0.0001$. The temporal stability of all measures was much lower, although statistically significant in all but one instance. The kappa values ranged between 0.19 and 0.59. Table 1c presents the findings on temporal stability across measures, as well as across interviewers, for both 9 month and 18 month time spans. Interestingly, the kappa values were generally as high, or higher, than those within

measure. It is also noteworthy that, as for the within-measure across-time comparisons (Table 1b), the kappa values when different interviewers were used tended to be higher than when it was the same interviewer on the two occasions. Of course, the time span was greater (18 months rather than 9 months) when the interviewers were the same, so this probably accounted for the difference. It is also evident that the stability tended to be greater between T2 and T3 than between T1 and T2.

Table 2 presents the comparable findings for warmth. The agreements between the brief PACE-EE and the whole interview ratings for warmth were very high (ranging from 0.82 to 0.88 across the three interview occasions). The agreements between both of these and the number of positive comments tended to be somewhat lower (0.45 to 0.75). As for criticism, the temporal stability kappa values within each measure were mainly around 0.3 (range from 0.11 to 0.39). The values were closely similar when considered across measures as well as across interviewer and over time (range from 0.09 to 0.38). It is

Table 1. Agreements (kappa values) among measures and temporal stability of criticism

a) Within time, across measure	Time 1	Time 2	Time 3
PACE-EE vs WI-EE	0.88	0.97	0.94
PACE-EE vs no. crit. comm.	0.74	0.84	0.87
No. crit. comm. vs WI-EE	0.82	0.93	0.94
b) Within measure, across time	Different interviewers		Same interviewer
	T1/T2	T2/T3	T1/T3
PACE-EE	0.33**	0.59	0.35
WI-EE	0.33**	0.52	0.33**
No. crit. comm.	0.38	0.49	0.19 ns
c) Across measure, across time	Different interviewers		Same interviewer
	T1/T2	T2/T3	T1/T3
PACE-EE vs WI-EE	0.31**	0.55	0.31**
PACE-EE vs no. crit. comm.	0.28**	0.53	0.29**
WI-EE vs no. crit. comm.	0.31**	0.50	0.31**
WI-EE vs PACE-EE	0.35	0.56	0.37
No. crit. comm. vs WI-EE	0.40	0.51	0.21§
No. crit. comm. vs PACE-EE	0.42	0.55	0.25*

Significance: $p \leq 0.001$ unless separately indicated; ** $p \leq 0.01$; * $p \leq 0.05$; § $p = 0.051$; ns = not significant

N.B. 'No criticism and some criticism' and 'high criticism and very high criticism' were combined throughout, as otherwise too many empty, or nearly empty, cells impaired calculation of kappa.

Table 2. Agreements (kappa values) among measures and temporal stability of warmth

a) Within time, across measure	Time 1	Time 2	Time 3
PACE-EE vs WI-EE	0.88	0.82	0.86
PACE-EE vs no. pos. comm.	0.65	0.68	0.45
No pos. comm. vs WI-EE	0.65	0.75	0.56
b) Within measure, across time	Different interviewers		Same interviewer
	T1/T2	T2/T3	T1/T3
PACE-EE	0.33	0.29	0.26
WI-EE	0.39	0.30	0.32
No. pos. comm.	0.31	0.11 ns	0.25
c) Across measure, across time	Different interviewers		Same interviewer
	T1/T2	T2/T3	T1/T3
PACE-EE vs WI-EE	0.31	0.31	0.24**
PACE-EE vs no. pos. comm.	0.28	0.24	0.09 ns
WI-EE vs no. pos. comm.	0.33	0.20**	0.14 ns
WI-EE vs PACE-EE	0.38	0.26	0.31
No. pos. comm. vs WI-EE	0.30	0.22**	0.31
No. pos. comm. vs PACE-EE	0.27	0.24	0.26

Significance $p \leq 0.001$ unless separately indicated; ** $p \leq 0.01$; ns not significant. three-point scales used throughout.

notable that the lowest kappa values were ones that involved positive comments, which have a somewhat different focus from warmth (being based solely on the words used).

In many studies, attention is focused on categorical distinctions based on low warmth or high criticism. Accordingly, similar comparisons were made for such categories with results that were almost identical. Thus, of the 87 children, 22 were the target of high criticism at Time 1 on either the brief PACE-EE or the whole interview; of these 20 were above threshold on both measures. The comparable figures for Time 2 were 18 out of 19, and for Time 3 22 out of 23. For warmth, the figures were 20 out of 25 at Time 1, 19 out of 23 at Time 2, and 20 out of 25 at Time 3.

One of the key strengths of the kappa statistic is that it takes into account the probability of chance agreement as determined by base rates. Nevertheless, the interpretation of kappa values is necessarily influenced by the distributional characteristics of the ratings. The distribution of criticism ratings on the PACE-EE at baseline was: low = 17, some = 50, high

= 17, very high = 3. The distribution on the WI-EE at baseline was very similar: 13, 52, 19, and 3 respectively. The paucity of very high criticism ratings meant that 'high' and 'very high' needed to be combined (see Noda, Kraemer, Yesavage and Periyakoil, 2001, regarding need to avoid cells with small numbers). Possibly, 'low' and 'some' could have been kept separate, but in order to keep the cells with small numbers at a minimum, they were combined. The distribution of warmth ratings showed that a three-point scale was justified. At the baseline measurement, the distribution on the PACE-EE was: high = 26, medium = 37, and low = 24. The comparable figures for WI-EE were 27, 39 and 21 respectively. Accordingly, it is evident that the sample used included an adequate number of individuals with high/very high criticism and low/very low warmth, the extremes of particular clinical interest.

Focusing on those negative extremes, the proportion of cases with agreement within-time, across-measures, was 83.3% for PACE-EE versus WI-EE, corresponding to a kappa value of 0.88. The percentage

agreement for other comparisons in Table 1a were closely similar. The kappa values within measures and across time, and across measures, across time in Tables 1b and 1c were lower, with corresponding lower agreement on the negative extremes. Thus, for PACE-EE versus WI-EE between T1 and T2 the agreement was 45% and between T2 and T3 was 50%, in both cases in relation to a kappa value of 0.31. The findings for other cells in the table were closely similar. The distributions of all 18 measures used in Tables 1 and 2 (three measures across three time points for both warmth and criticism) are provided in Appendix 2.

Discussion

The first finding is that there was very high agreement between the brief PACE-EE ratings and the whole interview ratings for both criticism and warmth. This is what has been previously found for the 5-minute speech sample and it appears that these brief assessments work remarkably well. In this study, the PACE-EE ratings and the whole-interview ratings were made by the same researcher and it could be suggested that high level of agreement stemmed from a halo effect. In order to consider the likelihood that such a bias operated, we need to turn to the temporal stability findings.

As noted in the methods section, the Time 1/Time 2 association will always have involved different interviewers, whereas the interviewer will have been the same for Time 1/Time 3. The agreements were broadly similar in the two cases. Thus, for criticism the comparison is between a range of 0.33 to 0.59 with different interviewers and a range of 0.19 to 0.35 for the same interviewers. For warmth, the comparable ranges were 0.11 to 0.39 and 0.25 to 0.32. The same applies to the comparison between Time 2/Time 3 associations, which involved different interviewers, and the Time 1/Time 3 associations, which did not. The findings provide no suggestion that it made any appreciable difference if the interviewer was the same or different, although this conclusion must carry the caution that the time span was greater for the same interviewer comparisons. The same applied to the findings across measures (see Tables 1c and 2c).

The second set of findings concerned the moderate stability of the EE measures across time spans of 9 and 18 months. If the EE measures simply reflected a pervasive enduring personality trait of the

parent, very high stability would be expected and that was not found. It could be suggested that this may have been a consequence of only moderate reliability of the measures but this does not seem likely because previous studies (Hodes et al., 1999b; Peris and Baker, 2000; Scott and Campbell, 2001) have been consistent in showing high reliability (so much so that it was decided that it was not necessary to repeat the reliability assessments in this study). Accordingly, we may reject the possibility that EE is no more than a personality trait.

An opposite concern is that EE might represent just an ephemeral emotional attitude of no enduring significance. If that were the case, a near zero correlation would be expected and, again, that is not what was found. To the contrary, kappa values of around 0.3 to 0.4 were found for agreement across measures, and across interviewers, over time periods of both 9 months and 18 months. Clearly, although changes over time must have been involved, as would be expected from the construct, substantial persistence was evident.

The findings, therefore, provide considerable support for the underlying assumptions of the measure. However, it is important, too, to note what remains to be established. The EE construct was well ahead of its time in recognizing the importance of individual-specific emotions, a recognition that came before the later appreciation of this point that stemmed from behaviour genetic evidence (Plomin and Daniels, 1987). Nevertheless, there is a lack of evidence on both the extent of differences in parental EE as directed towards different children in the same family, and on the psychopathological risk implications of such differences. Moreover, little is known about the extent to which variations in EE are associated with differences in the parental treatment of different children in the same family – either as reported by parents or by the young people themselves. Finally, longitudinal data are needed in order to determine the extent to which changes in EE over time are driven by changes in child behaviour, or by changes in family relationships, or by changes in the individual whose EE is being measured.

The findings on EE show the utility of measures based on just very brief periods of non-directive interviewing, making them very practical in a wide range of studies in which EE is just one of a larger set

of measures. The need now is to use EE as a way of exploring the development and consequences of child-specific differences in negative parental emotions.

References

- Asarnow JR, Goldstein MJ, Thompson M, Guthrie D. One-year outcomes of depressive disorders in child psychiatric inpatients: evaluation of the prognostic power of a brief measure of expressed emotion. *J Child Psychol Psychiatry* 1993; 34: 129–37.
- Asarnow JR, Thompson M, Hamilton EB, Goldstein MJ, Guthrie D. Family-expressed emotion, childhood-onset depression, and childhood-onset schizophrenia spectrum disorders: Is expressed emotion a non-specific correlate of childhood psychopathology or a specific risk factor for depression? *J Abnorm Child Psychol* 1994; 22: 129–46.
- Baker BL, Heller TL, Henker B. Expressed emotion, parenting stress, and adjustment in mothers of young children with behavior problems. Unpublished manuscript, Department of Psychology, UCLA, 1998.
- Baker BL, Heller TL, Henker B. Expressed emotion, parenting stress, and adjustment in mothers of young children with behavior problems. *J Child Psychol Psychiatry* 2000; 41: 907–15.
- Barnes-McGuire J, Earls F. Research note: the test-retest stability of the Five Minute Speech Sample in parents of disadvantaged minority children. *J Child Psychol Psychiatry* 1994; 35: 971–9.
- Brown GW, Birley JLT, Wing JK. Influence of family life on the course of schizophrenic disorders: a replication. *Br J Psychiatry* 1972; 121: 241–58.
- Brown GW, Monck EM, Wing JK. Influence of family life on the course of schizophrenic illness. *J Prev Soc Med* 1962; 16: 55–68.
- Brown GW, Rutter ML. The measurement of family activities and relationships: a methodological study. *Human Relations* 1966; 19: 241–63.
- Cohen, J. Weighted Kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Med* 1968; 70: 213–19.
- Glen S, Simpson A, Drinnan D, McGuinness D, Sandberg S. Testing the reliability of a new measure of life events and experiences in childhood: The Psychosocial Assessment of Childhood Experiences (PACE). *Eur Child Adolesc Psychiatry* 1993; 2: 98–110.
- Hahlweg K, Goldstein MJ, Nuechterlein K, Magana A, Mintz J, Doane J, Miklowitz D, Snyder K. Expressed emotion and patient-relative interaction in families of recent onset schizophrenics. *J Consult Clin Psychol* 1989; 57: 11–18.
- Hibbs ED, Hamburger SD, Lenane M, Rapoport JL, Kruesi MJP, Keysor CS, Goldstein MJ. Determinants of expressed emotion in families of disturbed and normal children. *J Child Psychol Psychiatry* 1991; 32: 757–70.
- Hibbs ED, Hamburger SD, Kruesi MJP, Lenane M. Factors affecting expressed emotion in parents of ill and normal children. *Am J Orthopsychiatry* 1993; 63: 103–12.
- Hirshfeld DR, Biederman J, Brody L, Faraone SV, Rosenbaum JF. Associations between expressed emotion and child behavioral inhibition and psychopathology: a pilot study. *J Am Acad Child Adolesc Psychiatry* 1997; 36: 205–13.
- Hodes M, Dare C, Dodge E, Eisler I. The assessment of expressed emotion in a standardised family interview. *J Child Psychol Psychiatry* 1999a; 40: 617–25.
- Hodes M, Garralda ME, Rose G, Schwartz R. Maternal expressed emotion and adjustment in children with epilepsy. *J Child Psychol Psychiatry* 1999b; 40: 1083–93.
- Kavanagh DJ. Recent developments in expressed emotion and schizophrenia. *Br J Psychiatry* 1992; 160: 601–20.
- Kershner JG, Cohen NJ, Coyne JC. Expressed emotion in families of clinically referred and non-referred children: toward a further understanding of the expressed emotion index. *J Fam Psychol* 1996; 10: 97–106.
- King S. Is expressed emotion cause or effect? A longitudinal study. *Int Clin Psychopharmacol* 1998; 13: 107–8.
- Kraemer HC, Periyakoil VS, Noda A. Tutorial in biostatistics: kappa coefficients in medical research. *Stat Med* 2002; 21: 2109–29.
- Liakopoulou M, Alifireraki T, Katideniou A, Peppas M, Maniati M, Tzikas D, Hibbs ED, Dacou-Voutetakis C. Maternal expressed emotion and metabolic control of children and adolescents with diabetes mellitus. *Psychother Psychosom* 2001; 70: 78–85.
- Leeb B, Hahlweg K, Goldstein MJ, Feinstein E, Mueller U, Dose M, Magana-Amato A. Cross-national reliability, concurrent validity, and stability of a brief method for assessing expressed emotion. *Psychiatry Res* 1991; 39: 25–31.
- Leff J, Vaughn CE. *Expressed Emotion in Families: Its Significance for Mental Illness*. New York: Guilford Press, 1985.
- McCreadie RG, Robertson LJ, Hall DJ, Berry I. The Nithsdale schizophrenia surveys: XI. Relatives' expressed emotion: stability over five years and relation to relapse. *Br J Psychiatry* 1993; 162: 393–7.
- McGuire JB, Earls F. The test-retest stability of the five minute speech sample in parents of disadvantaged minority children. *J Child Psychol Psychiatry* 1994; 35: 38–41.
- Magaña AB, Goldstein MJ, Karno M, Miklowitz DJ, Jenkins J, Falloon IRH. A brief method for assessing expressed emotion in relatives of psychiatric patients. *Psychiatry Res* 1986; 17: 203–12.
- Malla A, Kazarian S, Barnes S, Cole J. Validation of the five-minute speech sample in measuring expressed emotion. *Can J Psychiatry* 1991; 36: 297–9.

- Miklowitz DJ, Goldstein MJ, Falloon IRH, Doane JA. Interactional correlates of expressed emotion in families of schizophrenics. *Br J Psychiatry* 1984; 144: 482–7.
- Noda AM, Kraemer HC, Yesavage JA, Periyakoil VS. How many raters are needed for a reliable diagnosis? *Int J of Methods in Psychiatric Res* 2001; 10: 119–23.
- Nugter MA, Dingemans PM, Linszen AJ, Van der Does AJW, Gerson B. The relationship between expressed emotion, affective style and communication deviance in recent-onset schizophrenia. *Acta Psychiatr Scand* 1997; 96: 445–51.
- Patterson P, Birchwood M, Cochrane R. Preventing the entrenchment of high expressed emotion in first episode psychosis: early developmental attachment pathways. *Aust N Z J Psychiatry* 2000; 34: S191–7.
- Peris TS, Baker BL. Applications of the expressed emotion concept to young children with externalising behavior: Stability and prediction over time. *J Child Psychol Psychiatry* 2000; 41: 457–62.
- Plomin R, Daniels D. Why are children in the same family so different from one another? *Behav Brain Sci* 1987; 10: 1–15.
- Rutter M. *Children of Sick Parents: An Environmental and Psychiatric Study*. Institute of Psychiatry Maudsley Monographs No.16. London: Oxford University Press 1966.
- Rutter M, Brown GW. The reliability and validity of measures of family life and relationships in families containing a psychiatric patient. *Social Psychiatry* 1966; 1: 38–53.
- Rutter M, Quinton, D. Parental psychiatric disorder: effects on children. *Psychol Med* 1984; 14: 853–80.
- Sandberg S, Paton JY, Ahola S, McCann DC, McGuinness D, Hillary CR, Oja H. The role of acute and chronic stress in asthma attacks in children. *Lancet* 2000; 356: 982–7.
- Sandberg S, Rutter M, Giles S, Owen A, Champion L, Nicholls J, Prior V, McGuinness D, Drinnan D. Assessment of psychosocial experiences in childhood: Methodological issues and some illustrative findings. *J Child Psychol Psychiatry* 1993; 34: 879–97.
- Scott S, Campbell C. Expressed emotion about children: reliability and validity of a Camberwell Family Interview for Childhood (CFI-C). *Int J of Methods in Psychiatric Res* 2001; 9: 3–10.
- Schreiber JL, Breier AL, Pickar D. Expressed emotion: trait or state? *Br J Psychiatry* 1995; 166: 647–9.
- Schwartz CE, Dorer DJ, Beardslee WR, Lavori PW, Keller MB. Maternal expressed emotion and parental affective disorder: risk for childhood depressive disorder, substance abuse, or conduct disorder. *J Psychiatric Res* 1990; 24: 231–50.
- Stubbe DE, Zahner GEP, Goldstein MJ, Leckman JF. Diagnostic specificity of a brief measure of expressed emotion: a community study of children. *J Child Psychol Psychiatry* 1993; 34: 139–54.
- Tarrier N, Sommerfield C, Pilgrim H. Relatives' expressed emotion (EE) and PTSD treatment outcome. *Psychol Med* 1999; 29: 801–18.
- Taylor E, Sandberg S, Thorley G, Giles S. *Epidemiological Study of Hyperactivity of Childhood*. Maudsley Monograph No 33. Oxford: Oxford University Press, 1991.
- Van Furth E, Van Strien D, Van Son M, Van Engeland H. The validity of the Five-Minute Speech Sample as an index of expressed emotion in parents of eating disorder patients. *J Child Psychol Psychiatry* 1993; 34: 1253–60.
- Vaughn C, Leff, J. The measurement of Expressed Emotion in families of psychiatric patients. *Br J Soc Clin Psychol* 1976; 15: 157–65.
- Vostanis P, Nicholls J. Nine-month changes of maternal expressed emotion in conduct and emotional disorders of childhood: a follow-up study. *J Child Psychol Psychiatry* 1995; 36: 833–46.
- Vostanis P, Nicholls J, Harrington R. Maternal expressed emotion in conduct and emotional disorders of childhood. *J Child Psychol Psychiatry* 1994; 35: 365–76.
- Wamboldt FS, O'Connor SL, Wamboldt MZ, Gavin LA, Klinnert MD. The Five Minute Speech Sample in children with asthma: Deconstructing the construct of expressed emotion. *J Child Psychol Psychiatry* 2000; 41: 887–98.
- Wamboldt FS, Wamboldt MZ, Gavin LA, Roesler TA, Brugman SM. Parental criticism and treatment outcome in adolescents hospitalised for severe chronic asthma. *J Psychosom Res* 1995; 39: 995–1005.

Correspondence: Seija Sandberg, Consultant and Senior Lecturer, Department of Psychiatry and Behavioural Sciences, Royal Free and University College Medical School, Wolfson Building, 48 Riding House Street, London W1N 8AA.
Telephone: (+44)01708 796 800 (clinic), (+44)077113 44717 (mobile).
Email: s.sandberg@ucl.ac.uk

Appendix 1: PACE – expressed emotion

What is he/she like as a person _____ ,
I mean as an individual?

(GIVE NON-DIRECTIVE ENCOURAGEMENT TO
HELP INFORMANT EXTEND DESCRIPTION)

How would you describe his/her personality?

How do you mean? In what way? Can you tell me a little
more about that?

Are there any ways that you would like him/her to be
different?

Would you say he/she is an easy or difficult person to be
friendly/affectionate with?

What would you pick out as his/her most difficult charac-
teristics?

What is his/her best feature, do you think?

Are you glad that she is a girl (he is a boy) or would you
have preferred a boy (girl)?

Appendix 2: distribution of variables

CRITICISM		Low/moderate	High/very high
PACE-EE criticism	Time 1	67	20
WI-EE criticism	Time 1	65	22
No critical comments	Time 1	65	22
PACE-EE criticism	Time 2	69	18
WI-EE criticism	Time 2	68	19
No. critical comments	Time 2	70	17
PACE-EE Criticism	Time 3	64	23
WI-EE Criticism	Time 3	65	22
No. critical comments	Time 3	67	20

WARMTH		High	Moderate	Low/very low
PACE-EE warmth	Time 1	26	37	24
WI-EE warmth	Time 1	27	39	21
No. positive comments	Time 1	25	37	25
PACE-EE warmth	Time 2	24	40	23
WI-EE warmth	Time 2	30	38	19
No. positive comments	Time 2	27	38	22
PACE-EE warmth	Time 3	26	36	25
WI-EE warmth	Time 3	29	38	20
No. positive comments	Time 3	18	42	27