

EDITORIAL

Stressor, vulnerability and depression: a question of replication¹

One of the most fertile areas in social psychiatry in the last two decades has been the study of the aetiology of depressive illness. The number of reports appearing requires periodic stocktaking to keep abreast of developments, but the criteria by which findings should be evaluated are not always self-evident and merit some consideration. This theme is particularly topical in view of a recent editorial in *Psychological Medicine* which, in discussing one theoretical model of depressive disorder in women, concludes decisively: 'In summary Brown & Harris' vulnerability model has not been replicated' (Tennant, 1985). In commenting on and disagreeing with this conclusion, we hope to illuminate a number of more general issues surrounding replication studies and the criteria for their interpretation.

There are various approaches to the study of life events, vulnerability and psychiatric disorder. Some 90% of all studies have employed questionnaire-based measures that have proved not only to be burdened with problems of low reliability and poor validity, but also to have a predilection for reflecting general rather than specific phenomena (e.g. cumulative life-event scores for a period of time rather than the characteristics such as threatfulness of particular life events). The problems arising from the use of such measures have been compounded rather than eased by the use of more complex research designs. For example, longitudinal studies have typically failed to utilize life events occurring in the follow-up period, presumably because of the problem with such measures of establishing time order in relation to any onset (e.g. Henderson *et al.* 1981). Since events of aetiological importance tend to occur within a few weeks of onset, many relevant stressors are bound to be missed.

An alternative tradition relying on intensive interviews and rating scales completed by investigators has, in contrast, developed reliable and apparently valid measures and established the presence of large associations between certain types of stressor and the onset of depression. Results on the whole have been so clear that the main focus of recent interest has been the possibility that certain ongoing vulnerability factors increase risk, but only in the presence of a stressor. It is important to distinguish here the replicability of the finding that a high proportion of depressive onsets are preceded by stressors of these types from the replicability, at a different level, of a vulnerability model. If conceived broadly, this model implies some form of synergy between stressor and vulnerability factor in which the risk of onset when a stressor is experienced without a vulnerability factor is considerably higher than when a vulnerability factor is experienced without a stressor, although this may exceed the risk when neither factor is present. In narrower terms, the vulnerability model implies that when a vulnerability factor is present without a stressor, the risk is no higher than in the absence of both. The Brown–Harris model has typically been defined in the narrower of these traditions, although the authors have stated that they do not see any difference of any particular theoretical significance between the two (Brown & Harris, 1978*b*), especially as measurement error will serve to reduce the chances of confirming the narrower version (Brown & Harris, 1986). It is therefore worth stating at the outset in discussing replication that too absolute a distinction between the broader and narrower versions at this stage of research may lose the wood for the trees. The message which both versions seek to convey is that the onset of depression is basically in response to a stressor, specifically characterized as a provoking agent (for definition, see Brown & Harris, 1978*a*), and that vulnerable subjects are considerably more likely to respond in this way. Either version has implications for the

¹ Address for correspondence: Professor George W. Brown, Department of Social Policy and Social Science, Royal Holloway and Bedford New College, 11 Bedford Square, London WC1B 3RA.

evolution of further hypotheses in that by focusing upon the stressor and its particular qualities (in this case, severe loss experiences) an understanding of depressive cognitions is suggested in terms of personal meaning and generalized hopelessness. In other words, it would be misleading to reject the vulnerability model as a whole just because it does not seem to be supported in the narrower sense. This is particularly important given the statistical niceties involved in testing for synergy, where there is still no general agreement over the choice between tests for additive and multiplicative interaction (Everitt & Smith, 1979; Parry & Shapiro, 1986; Brown & Harris, 1986; Brown, 1986).

In his editorial about replications of the Brown–Harris model, Tennant himself mentions some of these problems, and at one point he selects a small number of studies which have used measures comparable with those of Brown & Harris, although in 1 of these 6 studies (Solomon & Bromet, 1982) the LEDS measure of stressors was not used. Focusing upon the main vulnerability factor originally considered, a measure of intimacy, Tennant states at this point that ‘three studies found interaction between lack of intimacy, life events and psychological disorder...[which]...lends some support to one part of Brown & Harris’ vulnerability model’. But two of the studies he cites as non-replications have almost identical results – in terms of the size of the predicted trend – to those accepted as successful: they just failed to reach statistical significance with the test employed because of small numbers. In fact, there have been, excluding the 2 original Camberwell surveys, 8 reasonably strict replications of the original study, and all but the one by Bebbington and his colleagues (1984) obtained essentially comparable results to the original survey (Table 1). Two studies using related measures also obtained similar results (Paykel *et al.* 1980; Murphy, 1982). This leaves only 1 of the 8 studies (or 10 in the full series) that failed to replicate the original work, and this study measures one of the provoking agents, major difficulties, rather differently (Bebbington *et al.* 1984). It would seem, then, that with a wider range of appropriate studies than Tennant selects, and a more considered approach to the interpretation of levels of statistical significance, the degree of support for this model might be judged as high.

In the body of the review Tennant appears to agree with this, but perhaps it is the mass of other studies cited which intervenes to cause the dysjunction of his earlier statements and his final dismissive generalization. These twenty other studies come from such a different tradition of measurement that they are irrelevant when judged as possible replications. (There are many cogent reasons why they would be inefficient in demonstrating a vulnerability effect (Brown & Harris, 1986).)

There is an interesting dilemma here. Cook & Campbell (1979), in a general discussion of issues of validity, point out the advantage of using somewhat different definitions of basic concepts. Any replication would be that much more impressive, and in the longer term new insights are likely to emerge from the clash of perspectives. However, it does not follow that failure to replicate in such circumstances (that is, with differing measures) should be taken as an equally impressive disconfirmation. It is not always realized that due to the relatively small samples demanded by the use of investigator-based measures quite modest changes in procedure can have important effects. In discussing a recent replication of the work with expressed emotion (EE) and relapse in schizophrenia by a team at Northwick Park, Leff & Vaughn (1986) note that of the 60 patients living in parental households both parents were interviewed in only 6 instances. The convention of EE research has been to classify a household as high even if only one of the two parents is rated as high EE. The failure to interview one of the parents must lead to the incorrect classification of a number of households. Such insistence on strict replication of procedures does not rule out changes; it asserts that the original procedures should *also* be followed so that the effect of changes can be monitored. Misconceptions are bound to proliferate if this rule is not followed.

Another implication of research with small numbers is that results failing to reach statistical significance in terms of a traditional 5% level can still represent, as with the earlier examples, a confirmation. There is a great difference between reports where the results are of the same order as the original finding but just fall short of a two-tailed test of statistical significance, and those where there is no trend or it is reversed. A one-tailed test of statistical significance is fully admissible in the former circumstances.

Table 1. Onset of 'caseness' depression, intimacy with husband and presence of a severe event or major difficulty

	Provoking agent		No provoking agent		P
	No intimacy (% onset)	Intimacy (% onset)	No intimacy (% onset)	Intimacy (% onset)	
<i>Strict replication of Camberwell survey</i>					
(A) Severe events or major difficulty					
Brown & Harris (1978 ^a) (Camberwell)	32 (24/76)	10 (9/88)	3 (2/62)	1 (2/193)	NS
Campbell <i>et al.</i> (1983) (Oxford)	46 (12/26)	13 (3/23)	12 (2/17)	2 (1/44)	NS
Brown & Prudo (1981) (Lewis)	36 (8/22)	15 (5/33)	5 (2/40)	1 (1/92)	NS
Bebington <i>et al.</i> (1984) (Camberwell)	22 (8/37)	24 (5/21)	11 (5/45)	6 (3/49)	NS
Parry & Shapiro (1986) (Sheffield)	31 (8/26)	10 (5/49)	10 (2/19)	5 (5/98)	NS
Finlay-Jones ^a (Regents Park area, London)	45 (24/53)	17 (4/23)	8 (3/36)	3 (1/39)	NS
Brown & Andrews (1985) (Islington)	26 (23/90)	12 (7/60)	3 (1/38)	1 (1/115)	NS
(B) Severe events only					
Costello (1982) (Calgary)	57 (8/14)	21 (5/24)	7 (4/56)	5 (14/292)	NS
Martin (1985) ^b (Manchester)	73 (8/11)	14 (2/14)	40 (2/5)	4 (2/47)	P < 0.01
Total	35 (123/355)	13 (45/335)	7 (23/318)	3 (30/969)	—
<i>Closely related studies</i>					
Paykel <i>et al.</i> (1980) ^c (South London)	82 (9/11)	24 (8/34)	14 (1/7)	12 (6/52)	NS
Murphy (1982) ^d (North London)	35 (6/17)	11 (10/90)	0 (0/19)	3 (2/74)	NS
Overall total	36 (138/383)	14 (63/459)	7 (24/344)	3 (38/1095)	—

The values shown are percentages, with the numbers of subjects in parentheses.

^a R. A. Finlay-Jones (personal communication).

^b A study of a series of women covering a pregnancy and a birth in every instance.

^c Study of post-partum women: one or more undesirable events and poor communication with husband.

^d Elderly sample of both sexes: severe event or major difficulty. Low intimacy = no confiding with anyone, i.e. not only husband.

When judging replications of a model rather than of merely a simple finding, statistical interpretation has been further complicated by the introduction of complex procedures to analyse such results, especially those concerning vulnerability. There has been considerable resulting confusion over the question of statistical interaction. Earlier statistical conventions required the presence of interaction in no more than the 'additive' sense, as used, for example, in ordinary regression with categorical variables. Log-linear modelling used in a number of recent analyses requires 'multiplicative' interaction, and in this sense the original Camberwell results do not show interaction. However, it can be stated at once that it is misleading to cite such a failure on the part of new studies as an instance of non-replication. As we have already seen in terms of the patterning of results, they closely parallel the original findings. The second point is that the different statistics are simply different ways of describing the same result and make no difference to their practical implications. Table 1, in addition to confirming the presence of synergy, shows that the proportion of those developing depression who have a stressor *and* lack of intimacy is particularly high (half of all the onsets) and that relatively few onsets occur among those with *only* lack of intimacy (9% of all onsets). The failure to show statistical interaction using logistic regression does not detract from this clear conclusion. The matter is, in fact, a theoretical not a statistical issue. It must be settled by asking what it is theoretically relevant to require a statistic to represent. There is no inherent truth in a statistical technique. We have argued elsewhere that additive interaction is a sensible way in which to view such phenomena, but will not pursue the matter here as it is not strictly pertinent to the question of replication (Brown & Harris, 1986, pp. 141–157; Brown, 1986).

The phrase 'vulnerability model' is, of course, more all-embracing than the provoking agent-intimacy synergism: there were three other vulnerability factors mentioned in the original Camberwell report and perhaps the definitive tone of Tennant's conclusion derives as much from results involving the other three as from the studies of intimacy and social support which are discussed more extensively in the review. It may be useful here to raise a more general point before discussing these other three factors. In judging replications, almost as important as considerations of measures, samples, timing, grouping and power of statistical tests is clarity in specifying whether replication is being considered at the level of findings and related model, or of theory. Replication can occur at one but not the other level. Insofar as it is wished to seek confirmation at the theoretical level, failure to replicate at the level of results will not necessarily be damaging and replication at the level of results may in the longer term not be enough. This arises from the use, particularly in early stages of research, of measures that are at best only approximate to what is relevant in a theoretical sense. For example, in Camberwell in 1969 women with 3 or more children under the age of 15 at home were at a higher risk for depression in the sense of being more *vulnerable* once a stressor had occurred. This, in turn, was interpreted in terms of a theory that postulated that low self-esteem and helplessness were crucial for the aetiology of depression and that having 3 or more children was associated with these factors. Two possible directions of causality were suggested: on the one hand, such children by 'trapping' women more in the home might render them more helpless in a crisis than those with fewer children; and, on the other hand, the more helpless women might well have been more likely to have more than two children in the first place, as a result of their failure to master contraceptives and demanding husbands. However, in a recent survey in Islington in 1980 those with 3 or more children were no more at risk of depression than were other women with children at home. At the level of measurement this represents a clear non-replication but at the level of theory it might well prove to be supportive. It is possible that easier contraception (especially the IUD and terminations), and increases in the availability of nursery school places, had led in the 1970s to important changes in the kind of women with 3 or more children. In Islington we found that women who had had a third child before 1970 were more helpless and had lower self-esteem than either those women completing their families of one and two before 1970 or those with a third child between 1971 and 1980 (for a full exposition, see Harris *et al.* 1986*b*). In other words, with changing social contexts a demographic-type variable of this kind may relate to a quite different underlying theoretical concept. Given this, failure to replicate may not prove to be a disconfirmation of the underlying theory. We use this example as an illustration and do not wish to place any particular

weight on it. However, it is sufficient to underline the need to move, as new studies are carried out, to more sensitive and theory-based measures. Demographic-type variables, such as number of children, are in particular danger of non-replication in terms of changes in milieu over time or space.

At this point it is possible to return to the other three vulnerability factors of the original model. One, lack of employment outside the home, was originally only acknowledged as a *second-order* factor once intimacy was taken into account (Brown & Harris, 1978a, p. 181). Three or more children has been replicated in two studies (Brown & Prudo, 1981; Campbell *et al.* 1983) but, as just discussed, this may be an example of a case where a failure to replicate ultimately provides support for an overall theoretical model. Research in terms of the findings concerning childhood loss of mother has had a patchy history, but this has probably been due at least in part to a failure to adopt the same measure as the Brown–Harris model. The variable was defined there as death or *continuous separation for at least one year* before the age of 11. Many studies report too low a rate of childhood loss of mother in their samples for analysis to be undertaken, but subsequent personal communication reveals that investigators had asked systematically only about deaths, not permanent separations. Even in studies where all separations from mother and their durations were collected, a proper replication of the Brown–Harris model has not always been undertaken (e.g. Bebbington *et al.* 1984). A high rate of depression among those with early loss has, in fact, been found in a study in Walthamstow specially designed to follow up the Camberwell finding (Brown *et al.* 1986a; Harris *et al.* 1986a). However, this work and also other more recent research has suggested that loss of mother itself may well be only an indicator of a more crucial factor for depression, namely poor parental care (Parker, 1981, 1983; Harris *et al.* 1986a; Bifulco *et al.* 1986; Kennard & Birchnell, 1982; Birchnell & Kennard, 1984). Future studies would therefore do best to concentrate attention on the quality of care following early loss experiences rather than the loss itself as a vulnerability factor, once again moving from the level of specific finding to that of theoretical interpretation. It is worth spelling out here why. Poor care from parent, or parent surrogate, like poor intimacy in an adult relationship, can be expected to be closely associated with a poor self-image. Because the regard of others is often a source of one's own self-regard, those exposed to deficient care in childhood are likely to place a lower estimate on their own worth, even perhaps when childhood is past, and will thus prove more vulnerable to depression in the face of provoking agents. Since low self-esteem was the intrapsychic lynchpin of the theoretical model developed to account for the Brown–Harris vulnerability findings, any results supporting this perspective deserve particularly serious consideration. It may not therefore be inappropriate to end by citing a set of new findings which bear on the model discussed here, since they involve measures of the lynchpin variable of self-esteem.

In a longitudinal enquiry in Islington, 150 women experienced a provoking agent in the follow-up year and for these the onset was nearly three times more likely among those with low self-esteem at first interview (Brown & Bifulco, 1985; Brown *et al.* 1986b). Low self-esteem was also associated with negative interaction in marriage (Brown & Andrews, 1986) and poor parental care in childhood (Bifulco *et al.* 1986), and both of these other factors also increased the risk of depression after a provoking agent in the manner predicted of vulnerability factors (namely, without a provoking agent the risk of onset was no higher than among those without low self-esteem, negative marital interaction or lack of care). A parallel study in Edinburgh has not yet reported on onset during the follow-up period, but found similar associations at first interview between low self-esteem (measured by the Rosenberg scale) and both low intimacy and childhood loss, even when cases with psychiatric disorder were excluded from the comparison (Ingham *et al.* 1986). Thus, in the light of the findings reviewed, it is perhaps wiser to conclude, unlike Tennant, that the Brown–Harris model has obtained a degree of support from a wide range of findings.

GEORGE W. BROWN AND TIRRIL HARRIS

REFERENCES

- Bebbington, P. E., Sturt, E., Tennant, C. & Hurry, J. (1984). Misfortune and resilience: a community study of women. *Psychological Medicine* **14**, 347–364.
- Bifulco, A., Brown, G. W. & Harris, T. O. (1986). Loss of parent, lack of parental care and adult psychiatric disorder: the Islington study. (MS)

- Birtchnell, J. & Kennard, J. (1984). Early and current factors associated with poor-quality marriage. *Social Psychiatry* **19**, 31–40.
- Brown, G. W. (1986). Statistical interaction and the role of social factors in the aetiology of clinical depression. *Sociology* (in the press).
- Brown, G. W. & Andrews, B. (1985). Comparison of Camberwell and Islington intimacy rating. (MS)
- Brown, G. W. & Andrews, B. (1986). Social support and depression. In *Dynamics of Stress* (ed. M. H. Appley and R. Trumbull). Plenum: New York.
- Brown, G. W. & Bifulco, A. (1985). Social support, life events and depression. In *Social Support: Theory, Research and Applications* (ed. I. Sarason). Martinus Nijhoff: Dordrecht.
- Brown, G. W. & Harris, T. O. (1978a). *Social Origins of Depression: A Study of Psychiatric Disorder in Women*. Tavistock Publications: London.
- Brown, G. W. & Harris, T. O. (1978b). Social origins of depression: a reply. *Psychological Medicine* **8**, 577–588.
- Brown, G. W. & Harris, T. O. (1986). Establishing causal links: the Bedford College Studies of Depression. In *Life Events and Psychiatric Disorders* (ed. H. Katschnig). Cambridge University Press: Cambridge.
- Brown, G. W. & Prudo, R. (1981). Psychiatric disorder in a rural and an urban population I: Aetiology of depression. *Psychological Medicine* **11**, 581–599.
- Brown, G. W., Harris, T. O. & Bifulco, A. (1986a). Long-term effect of early loss of parent. In *Depression of Childhood: Developmental Perspectives* (ed. M. Rutter, C. Izard and P. Read). Guilford Press: New York.
- Brown, G. W., Andrews, B., Harris, T., Adler, Z. & Bridge, L. (1986b). Social support, self-esteem and depression. *Psychological Medicine* (in the press).
- Campbell, E., Cope, S. & Teasdale, J. (1983). Social factors and affective disorder: an investigation of Brown and Harris's model. *British Journal of Psychiatry* **143**, 548–553.
- Cook, T. D. & Campbell, D. T. (1979). *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Houghton Mifflin: Boston.
- Costello, C. G. (1982). Social factors associated with depression: a retrospective community study. *Psychological Medicine* **12**, 329–339.
- Everitt, B. S. & Smith, A. M. (1979). Interaction in contingency tables: a brief discussion of alternative definitions. *Psychological Medicine* **9**, 581–583.
- Harris, T. O., Brown, G. W. & Bifulco, A. (1986a). Loss of parent in childhood and adult psychiatric disorder: the role of lack of adequate parental care. *Psychological Medicine* (in the press).
- Harris, T. O., Adler, Z., Bridge, L. & Brown, G. W. (1986b). Instability of indicator variables and replication studies: depression in women and number of children at home. (MS)
- Henderson, S., Byrne, D. G. & Duncan-Jones, P. (1981). *Neurosis and the Social Environment*. Academic Press: London.
- Ingham, J. G., Kreitman, N. B., Miller, P. McC., Sashidharan, S. P. & Surtees, P. G. (1986). Self-esteem, vulnerability and psychiatric disorder in the community. *British Journal of Psychiatry* (in the press).
- Kennard, J. & Birtchnell, J. (1982). The mental health of early mother separated women. *Acta Psychiatrica Scandinavica* **65**, 388–402.
- Leff, J. & Vaughn, C. (1986). First episodes of schizophrenia. *British Journal of Psychiatry* **148**, 215.
- Martin, C. J. (1985). Stress in the puerperium. Ph.D. dissertation: University of Manchester.
- Murphy, E. (1982). Social origins of depression in old age. *British Journal of Psychiatry* **141**, 135–142.
- Parker, G. (1981). Parental reports of depressives: an investigation of several explanations. *Journal of Affective Disorders* **3**, 131–140.
- Parker, G. (1983). Parental 'affectionless control' as an antecedent to adult depression. *Archives of General Psychiatry* **40**, 956–960.
- Parry, G. & Shapiro, D. A. (1986). Life events and social support in working-class mothers: stress-buffering or independent effects? *Archives of General Psychiatry* **43**, 315–323.
- Parry, G. & Shapiro, D. & Davies, L. (1981) Reliability of life-event ratings: an independent replication. *British Journal of Clinical Psychology* **739**, 213–220.
- Paykel, E. S., Emms, E. M., Fletcher, J. & Rassaby, E. S. (1980). Life events and social support in puerperal depression. *British Journal of Psychiatry* **136**, 339–346.
- Solomon, Z. & Bromet, E. (1982). The role of social factors in affective disorder: an assessment of the vulnerability model of Brown and his colleagues. *Psychological Medicine* **12**, 123–130.
- Tennant, C. (1985). Female vulnerability to depression. *Psychological Medicine* **15**, 733–737.