Expressed Emotion and Psychiatric Relapse

A Meta-analysis

Ronald L. Butzlaff, AM; Jill M. Hooley, DPhil

Background: Expressed emotion (EE) is a measure of the family environment that has been demonstrated to be a reliable psychosocial predictor of relapse in schizophrenia. However, in recent years some prominent nonreplications of the EE-relapse relationship have been published. To more fully address the question of the predictive validity of EE, we conducted a meta-analysis of all available EE and outcome studies in schizophrenia. We also examined the predictive validity of the EE construct for mood disorders and eating disorders.

Methods: An extensive literature search revealed 27 studies of the EE-outcome relationship in schizophrenia. Using meta-analytic procedures, we combined the findings of these investigations to provide an estimate of the effect size associated with the EE-relapse relationship. We also used meta-analysis to provide estimates of the effect sizes associated with EE for mood and eating disorders.

Results: The results confirmed that EE is a significant and robust predictor of relapse in schizophrenia. Additional analyses demonstrated that the EE-relapse relationship was strongest for patients with more chronic schizophrenic illness. Interestingly, although the EE construct is most closely associated with research in schizophrenia, the mean effect sizes for EE for both mood disorders and eating disorders were significantly higher than the mean effect size for schizophrenia.

Conclusion: These findings highlight the importance of EE in the understanding and prevention of relapse in a broad range of psychopathological conditions.

Arch Gen Psychiatry. 1998;55:547-552

Expessed emotion (EE) is a measure of the family environment that is based on how the relatives of a psychiatric patient spontaneously talk about the patient. Assessed during the Camberwell Family Interview (CFI), relatives are classified as being high in EE if they make more than a specified threshold number of critical comments or show any signs of hostility or marked emotional overinvolvement.

In the last 15 years, the EE construct has been extensively studied. More than 20 studies, conducted in many countries, have investigated the EE-relapse relationship in patients with schizophrenia. In addition, there is a growing literature concerning the role of EE in unipolar depression, bipolar disorder, and eating disorders. Expressed emotion has also been used in outcome studies of patients with dementia and diabetes mellitus. The results of these investigations make 2 things clear. First, rather than being a construct of interest solely with respect to schizophrenia, EE is a more general predictor of poor outcome across a range of conditions. Second, EE is a construct that is modifiable. Results from several trials of family-based treatment indicate that when family EE levels decrease, patients’ relapse rates also fall. From a clinical perspective, these findings are clearly very encouraging.

Given this, it is surprising that EE remains a somewhat controversial construct. However, based on the results of a small number of nonreplications, some clinicians and researchers have been quick to conclude that EE is not a reliable predictor of relapse. This article represents an effort to examine this issue in a statistically rigorous manner. Although aggregate analyses of the EE literature do exist, we chose to meta-analyze the studies because of the dangers of aggregating or pooling raw data without blocking, especially when using 2×2 tables. Moreover, because meta-analysis provides a way to combine similar studies in a manner that
MATERIALS AND METHODS

LITERATURE SEARCH

To be included, articles had to meet the following criteria: (1) patient diagnosis of schizophrenia or schizoaffective disorder or mood or eating disorder; (2) EE assessed using the CFI administered at the time of the index hospitalization (one exception was published before the CFI became the standard method for assessing EE); (3) EE used to predict relapse for 9 to 12 months; and (4) published data allowed an estimate of the effect size and significance level to be calculated. A total of 27 articles met these criteria. We excluded 22 experimental articles for the following reasons: (1) the CFI was not administered at intake; (2) the CFI was used to measure EE in nonfamily members (eg, nursing staff); (3) EE was used to predict something other than adult psychiatric relapse; (4) the sample subjects did not include both low- and high-EE families or the relapse data were not reported for both groups; and (5) the report described data that were also published elsewhere.

STATISTICAL ANALYSIS

The studies’ data were cast into 2 × 2 tables of counts—high vs low EE by relapsed vs not relapsed status. We chose to use *φ* as our measure of effect size because of the problems associated with other indices. In cases where authors did not report the number of subjects relapsing in the high- and low-EE groups, we used a reliable formula for calculating an effect size estimate. Three other studies reported relapse rates of 0 for one of the EE groups. We used a correction suggested by Overall in calculating the effect size estimates for these studies.

All effect size estimates were transformed into Fisher’s *z* scores before any other calculations were done to account for the nonnormal distribution of *r*. We also calculated the associated standard normal deviate *z* score for each study. This summary statistic is analogous to a *t* or *F* test in studies comparing differences between groups. Combining these *z* scores meta-analytically served as a test statistic for the estimate of the overall significance of the combined average effect size estimate of the studies. Table 1 details the studies used in this meta-analysis, noting the corrections described above.

LENGTH OF ILLNESS

In an earlier review, one of us (J.M.H.) suggested that there might be a relationship between the duration of schizophrenic illness and the magnitude of the EE-relapse link, with longer durations of illness being associated with greater effect sizes. We classified the EE reports according to the mean chronicity of the patient sample studied (Table 1).

Our categorization criteria were as follows. In the first group were studies in which the majority of patients were experiencing their first hospitalization. Our second category included studies with more heterogeneous samples, in which the mean chronicity of the patient sample was neither very recent nor very chronic (eg, a sample that contained 30% recent-onset patients but where the average number of prior hospitalizations was 2.8). The third category included studies with more chronic patients, where chronic was defined as more than 3 prior hospital admissions or a mean duration of illness of at least 5 years.

GEOGRAPHIC LOCATION

Recently, Bebbington and Kuipers used visual inspection of a graph to conclude that there was no variation in EE findings based on geographic location. We tested this hypothesis using an analysis of variance (ANOVA) method involving the Q statistic, which is much the same as a χ² statistic but using meta-analytic data. Following Bebbington and Kuipers, studies were grouped according to their most obvious broad geographic location (ie, Northern Europe, Southern Europe, North America, Australia, and Asia).

EE AND OTHER PSYCHIATRIC CONDITIONS

Expressed emotion was developed as a psychosocial predictor of relapse in schizophrenia. However, several researchers have documented the link between EE and relapse in patients with mood disorders and eating disorders, such as anorexia and obesity. We chose to meta-analyze these studies to establish the effect size of EE disorders other than schizophrenia. Because the number of studies in these areas is limited, readers should view these findings as preliminary.

EE AND MOOD DISORDERS

Six studies have examined the relationship between EE and relapse in patients with major mood disorders (Table 2). All found a positive association between EE and relapse. However, because of the relatively small number of studies that have been conducted, the number of critical comments required to classify families as high-EE has not been firmly established. Cutoff scores of 2 criticisms and 3 criticisms seem to have validity in unipolar patients. For patients with bipolar disorder, a cutoff score of 6 critical comments (ie, the cutoff score used for schizophrenia) is the most appropriate. Our results are based on cutoff scores of 2 and 3 for unipolar samples and 6 for bipolar samples. To facilitate future meta-analysis, we encourage researchers to report relapse by varying levels of critical comments, in addition to reporting the cutoff score that proves most significantly predictive.

RESULTS

HOW WELL DOES EE PREDICT RELAPSE IN SCHIZOPHRENIA?

The simple answer to this question is: quite well. All but 3 of the studies described in Table 1 (89%) showed a sig-

ARCH GEN PSYCHIATRY/VOL 55, JUNE 1998

©1998 American Medical Association. All rights reserved.
significant association between EE and patient relapse. If there were no relationship, we would expect 50% of the relapses that do occur and with two thirds of the relapses that do not occur.

How important are these findings?

It is not always easy to grasp the practical importance of a meta-analytic effect size $r = 0.30$. However, for a hypothetical sample of 200 patients (high EE = 100; low EE = 100), an effect size $r = 0.30$ translates into a high and low EE relapse rate of 65% and 35%, respectively. In this model, EE is associated with approximately one third of the patients resulted in a contrast $z = 1.93, P = .03$. As previously noted, EE appears to be a stronger predictor of relapse in patients with more longstanding illnesses.

Combining the effect sizes of individual studies in a meta-analysis requires that the assumption of homogeneity of variance be tested and met. We used the $\chi^2$ statistic to test for the heterogeneity of the effect sizes. The analyses resulted in a significant $\chi^2 = 54.01, df = 26, P = .001$. We therefore conducted additional analyses to identify the variable(s) that accounted for the heterogeneity. These are described below.

Length of Illness

The linear contrast of the relationship between effect size for the EE-relapse association and the chronicity category of the patients resulted in a contrast $z = 1.93, P = .03$.

WHAT ABOUT STUDIES THAT MIGHT HAVE BEEN OVERLOOKED?

The issue of unpublished research is often called the “file-drawer” problem. It is addressed using a calculation suggested by Rosenthal. To lower the significance level of this meta-analysis to just barely significant ($P = .05$), there would have to be 1246 new, unpublished, or undiscovered studies averaging null results. When we consider that research that does not confirm the relationship between EE and relapse often receives more attention than experimental replications, this seems unlikely.

Is Heterogeneity of Effect Sizes a Problem?

Combining the effect sizes of studies in the most chronic category still suggested by Rosenthal. The issue of unpublished research is often called the “file-drawer” problem. It is addressed using a calculation suggested by Rosenthal. To lower the significance level of this meta-analysis to just barely significant ($P = .05$), there would have to be 1246 new, unpublished, or undiscovered studies averaging null results. When we consider that research that does not confirm the relationship between EE and relapse often receives more attention than experimental replications, this seems unlikely.

Length of Illness

The linear contrast of the relationship between effect size for the EE-relapse association and the chronicity category of the patients resulted in a contrast $z = 1.93, P = .03$. As previously noted, EE appears to be a stronger predictor of relapse in patients with more longstanding illnesses.

Grouping studies according to patient chronicity accounted for the heterogeneity of the effect sizes for the recent-onset and the mixed chronicity groups. This suggests that illness chronicity was a hidden variable in the earlier combined analysis. Interestingly, however, the effect sizes of studies in the most chronic category still showed significant heterogeneity ($\chi^2 = 31.53, df = 12$, $P < .01$).

Author Information

**Table 1. Summary Table of Included Studies: Expresssed Emotion (EE) and Schizophrenia**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Low EE</th>
<th>High EE</th>
<th>$Zr$</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Araújo and Vazcarro²³</td>
<td>5/13</td>
<td>8/18</td>
<td>0.06</td>
<td>3</td>
</tr>
<tr>
<td>Barreto et al²²</td>
<td>0/12</td>
<td>8/24</td>
<td>0.35</td>
<td>1</td>
</tr>
<tr>
<td>Bertrand et al²²</td>
<td>4/18</td>
<td>14/24</td>
<td>0.38</td>
<td>3</td>
</tr>
<tr>
<td>Brown et al²⁴</td>
<td>13/47</td>
<td>38/50</td>
<td>0.53</td>
<td>3</td>
</tr>
<tr>
<td>Brown et al²³</td>
<td>9/56</td>
<td>26/45</td>
<td>0.47</td>
<td>3</td>
</tr>
<tr>
<td>Buekrem et al²⁵</td>
<td>...</td>
<td>...</td>
<td>0.17</td>
<td>2</td>
</tr>
<tr>
<td>Ito and Oshima²⁴</td>
<td>3/37</td>
<td>16/55</td>
<td>0.46</td>
<td>3</td>
</tr>
<tr>
<td>Ivanović and Vuletić²⁵</td>
<td>2/31</td>
<td>19/29</td>
<td>0.72</td>
<td>3</td>
</tr>
<tr>
<td>Kano et al²⁶</td>
<td>7/27</td>
<td>10/17</td>
<td>0.34</td>
<td>3</td>
</tr>
<tr>
<td>Köttgen et al²⁷</td>
<td>11/20</td>
<td>7/14</td>
<td>−0.05</td>
<td>2</td>
</tr>
<tr>
<td>Leff et al²⁹</td>
<td>5/54</td>
<td>5/16</td>
<td>0.27</td>
<td>1</td>
</tr>
<tr>
<td>Linszen et al²⁸</td>
<td>0/13</td>
<td>6/26</td>
<td>0.26</td>
<td>1</td>
</tr>
<tr>
<td>MacMillan et al²⁹</td>
<td>14/34</td>
<td>26/38</td>
<td>0.28</td>
<td>1</td>
</tr>
<tr>
<td>Molne et al³⁰</td>
<td>4/13</td>
<td>10/11</td>
<td>0.71</td>
<td>3</td>
</tr>
<tr>
<td>Montero et al³¹</td>
<td>7/31</td>
<td>8/28</td>
<td>0.07</td>
<td>2</td>
</tr>
<tr>
<td>Mönnig and Vojtěková³²</td>
<td>13/26</td>
<td>4/69</td>
<td>0.38</td>
<td>3</td>
</tr>
<tr>
<td>Niedermeier et al³³</td>
<td>6/21</td>
<td>16/28</td>
<td>0.29</td>
<td>2</td>
</tr>
<tr>
<td>Nuechterlein et al³⁴</td>
<td>0/12</td>
<td>12/31</td>
<td>0.37</td>
<td>1</td>
</tr>
<tr>
<td>Parker et al³⁴</td>
<td>9/15</td>
<td>20/42</td>
<td>−0.11</td>
<td>2-3</td>
</tr>
<tr>
<td>Phillips and Xiong³⁵</td>
<td>9/27</td>
<td>10/22</td>
<td>0.12</td>
<td>3</td>
</tr>
<tr>
<td>Rostovskovskova et al³⁶</td>
<td>1/15</td>
<td>11/25</td>
<td>0.51</td>
<td>...</td>
</tr>
<tr>
<td>Stirling et al³⁷</td>
<td>8/17</td>
<td>5/16</td>
<td>−0.16</td>
<td>1</td>
</tr>
<tr>
<td>Tanaka et al³⁸</td>
<td>6/28</td>
<td>14/24</td>
<td>0.40</td>
<td>3</td>
</tr>
<tr>
<td>Tarrer et al³⁹</td>
<td>4/19</td>
<td>14/29</td>
<td>0.28</td>
<td>2</td>
</tr>
<tr>
<td>Vaughan and Leff³²</td>
<td>1/16</td>
<td>10/21</td>
<td>0.48</td>
<td>2</td>
</tr>
<tr>
<td>Vaughan et al³³</td>
<td>3/18</td>
<td>20/66</td>
<td>0.39</td>
<td>3</td>
</tr>
<tr>
<td>Vaughan et al³⁴</td>
<td>10/41</td>
<td>25/47</td>
<td>0.30</td>
<td>3</td>
</tr>
</tbody>
</table>

*Ellipses indicate data not available.
†$Zr$ indicates z-to-r transformation.
‡Critical comments (CC) = 6 for bipolar samples.
§The sample also included 3 patients diagnosed with schizoaffective disorder.

**Table 2. Summary Table of Included Studies: Expresssed Emotion (EE) and Mood Disorders**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Low EE</th>
<th>High EE</th>
<th>$Zr$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hooley et al (unipolar)</td>
<td>0/5</td>
<td>20/34</td>
<td>0.37</td>
</tr>
<tr>
<td>Miklowitz et al (bipolar)</td>
<td>7/13</td>
<td>9/10</td>
<td>0.41</td>
</tr>
<tr>
<td>Miklowitz et al (bipolar/mania)</td>
<td>11/28</td>
<td>12/13</td>
<td>0.55</td>
</tr>
<tr>
<td>Okasha et al (unipolar/bipolar)</td>
<td>2/10</td>
<td>16/22</td>
<td>0.14</td>
</tr>
<tr>
<td>Priebe et al (bipolar)</td>
<td>3/10</td>
<td>9/11</td>
<td>0.58</td>
</tr>
<tr>
<td>Vaughan and Leff (unipolar)</td>
<td>2/9</td>
<td>14/21</td>
<td>0.44</td>
</tr>
</tbody>
</table>

*Zr indicates z-to-r transformation.
†Corrected using the method of Overall.
‡Critical comments (CC) = 6 for bipolar samples.
§The sample also included 3 patients diagnosed with schizoaffective disorder.

©1998 American Medical Association. All rights reserved.
Further examination revealed that this was largely attributable to Parker et al. They accounted for 12.94 of the $\chi^2$ value of 31.53. Methodological problems associated with this study may explain its outlier status.

**GEографIC LOCATION**

Does the magnitude of the EE-relapse association vary according to the location of the research site? The test of mean effect sizes across locations was not significant ($Q_{between} = 5.91, df = 4$). When categorized according to Bebbington and Kuipers, geographic location is not related to variations in effect size.

One advantage of this analytic approach was that it followed the grouping method described by Bebbington and Kuipers. Unfortunately, when we did this, the ANOVA assumption of homogeneity of variance was violated ($Q_{within} = 48.1$, $df = 22$, $P = .001$). Although ANOVA is relatively robust in this regard, we reanalyzed the data employing more narrow geographic boundaries. We assumed that accounting for more of the variance using location as a blocking variable would allow us to replicate the previously found nonsignificant effect of geographic region. However, this was not the case. Regrouping the studies as United Kingdom, Northern Europe, Southern Europe, Eastern Europe, North America, Australia, or Asia resulted in a significant variation in effect sizes between locations ($Q_{between} = 19.21$, $df = 6$, $P = .004$). Further examination revealed that the studies from Eastern Europe accounted for much of this heterogeneity (6.25 of the $Q_{between}$ statistic total). The second largest contributors to the total heterogeneity were the Australian studies, accounting for 4.46 of the total $Q_{between}$ statistic. These results suggest that the effect sizes of EE in Eastern Europe are unusually high. In contrast, the effect size associated with 1 of the 2 EE studies conducted in Australia is particularly low when compared with effect size estimates from other parts of the world.

**WHAT IS THE EFFECT SIZE FOR EE IN MOOD DISORDERS?**

With high EE defined as 2 or more critical comments, the mean and the weighted mean effect sizes were both $r = 0.39$. This effect size was associated with $z = 5.21$, $P < .001$. There was no significant heterogeneity in the effect sizes. When critical comments were set to 3 or more for high EE, the mean and weighted mean effect sizes rose to 0.45 ($z = 6.12$, $P < .001$), again with no significant heterogeneity in the data.

It is important to note that even when the EE cutoff for critical comments was set at 2 or more, we found a meta-analytic effect size $r = 0.39$. With a hypothetical group of 200 patients, this translates into expected relapse rates of 69.5% for patients with high-EE relatives and 30.5% for patients with low-EE relatives. The 95% CIs were also narrow (0.28-0.50). Our file drawer calculations show that there would need to be 54 unpublished studies to reduce the significance level to $P = .05$. When a critical comments cutoff of 3 or more was used, the 95% CIs narrowed even further (0.40-0.50) with a file-drawer statistic of 77. These findings provide strong support for the role of EE in the course of mood disorders.

**WHAT IS THE EFFECT SIZE FOR EE IN EATING DISORDERS?**

Three studies have reported on the relationship between EE and outcome in patients being treated for eating disorders. These studies differed from the schizophrenia studies in that outcome was measured by diet compliance, weight gain after treatment for obesity, and premature termination of treatment in patients with anorexia. The cutoffs for determining high EE also varied across the 3 studies. For our analysis, where more than 1 cutoff was provided we selected the cutoff score that would result in the smallest effect size.

All of the studies found a positive association between high EE and poor outcome. The weighted mean effect size was 0.51 ($z = 5.05$, $P < .001$), with 95% CIs of 0.36 to 0.70 and a file-drawer statistic of 25. As was the case for mood disorders, there was no significant heterogeneity of the effect sizes. Although only based on a few studies, our analyses suggest that EE is a strong predictor of early treatment outcome for patients with weight or eating disorders.

**IS EE A STRONGER PREDICTOR OF OUTCOME IN NONSCHIZOPHRENIA-RELATED CONDITIONS?**

The substantial and robust effect sizes for EE in mood and eating disorders raise the question of whether EE is a significantly better predictor of outcome when patients have disorders other than schizophrenia. We therefore compared the effect sizes associated with EE in schizophrenia, mood disorders, and eating disorders. The analyses revealed that EE was a significantly better predictor of outcome for mood disorders than it was for schizophrenia ($t_{[31]} = 1.93$, $P = .03$, using 3 critical comments as a cutoff for the depression studies). The same was also true for the comparison of the effect sizes of EE in schizophrenia vs eating disorders ($t_{[28]} = 2.03$, $P = .03$). The effect sizes for EE in mood disorders and eating disorders did not differ significantly, however. These findings suggest that although EE is a reliable predictor of poor outcome for schizophrenia, mood disorders, and eating disorders, EE is a significantly better predictor for the latter 2 disorders than it is for schizophrenia. Given the overwhelming amount of research that examines EE in families of patients with schizophrenia, this difference in effect sizes is an interesting and potentially important finding.

**COMMENT**

The results of this meta-analysis confirm that EE is a reliable predictor of relapse in patients with schizophrenia. Although several nonreplications exist, these do not require that the predictive validity of the construct be called into question.

The association between geographic location and the magnitude of the EE-relapse link is less clear. Although
our primary analysis revealed no significant relationship between study site and effect size, additional analysis indicated that this might not invariably be the case. Specifically, the effect size for EE appears to be unusually high in studies coming from Eastern Europe, and unusually low in one Australian investigation.24 More research from Eastern Europe would obviously be valuable, as would further examination of the EE-relapse link in Australian samples.

Although EE predicts relapse regardless of the chronicity of the patients being studied, the magnitude of the EE-relapse association increases when research samples contain more chronically ill patients. One explanation for this is that EE is a more reliable measure of the family environment when patients have been ill longer.30 Another possibility is that patients may become more sensitive to EE as the illness continues, perhaps through a process that resembles kindling or sensitization.31,32 This, of course, assumes that EE does indeed play some causal role in the relapse process. However, there is some evidence that this is the case.2,19

Schizophrenia is a disorder in which biological factors play a very central role. The role that psychosocial factors play might thus be somewhat restricted. That psychosocial factors are important is indicated by the significant EE-relapse relationship and by the success of family-based interventions designed to reduce patient relapse rates.33,34 However, EE may play an even more important role in the course of mood and eating disorders than schizophrenia. Examining the effect of family-based interventions on relapse rates and outcome for mood-disordered and eating-disordered patients is also likely to be a worthwhile avenue of empirical inquiry.

Finally, it warrants mention that more studies of the EE-relapse link in schizophrenia will not influence the effect sizes reported here in any appreciable way. Almost 40 years after the initial observation of Brown et al.,14 the elevated risk for relapse associated with high EE family environments appears well established. The time has now come for creative and sophisticated research that will tell us why EE is associated with relapse in such a wide range of psychopathological conditions.

Accepted for publication February 12, 1998.

We thank Robert Rosenthal, PhD, for statistical consultation and Lisa Vegge for assistance with the literature search.

Reprints: Jill M. Hooley, DPhil, Department of Psychology, Harvard University, 33 Kirkland St, Cambridge, MA 02138 (e-mail: jmh@wjh.harvard.edu).

REFERENCES

25. Ivanovic M, Vuletic Z. Expressed emotion in families of patients with frequent types of schizophrenia and influence on the course of illness: nine-months’ follow-up. Paper presented at: XIX Congress of the European Association of Behaviour Therapy; September 20, 1989; Vienna, Austria.

©1998 American Medical Association. All rights reserved.
36. Rostworowska M, Barbaro B, Cechnicki A. The influence of expressed emotion on the course of schizophrenia: a Polish replication. Poster presented at: 17th Congress of the European Association for Behavior Therapy; August 26, 1987; Amsterdam, the Netherlands.