Effectiveness of Activity-Based Group Work in Community Mental Health: A Systematic Review

Alison Bullock, Katrina Bannigan

KEY WORDS
- community mental health services
- evidence-based practice
- human activities
- mentally ill persons
- psychotherapy, group
- treatment outcome

Activity-based group work is widely used by occupational therapists in mental health settings (Lloyd, King, & Bassett, 2002), but the evidence to support this intervention is unclear. We therefore conducted a systematic review focused on the question, “Is activity-based group work effective in helping people with severe and enduring mental illness in community settings improve their functional ability and/or reduce their mental health symptoms?” We used a wide-ranging search strategy, including electronic searching, hand searching, citation searching, and use of gray literature, and identified 136 potentially relevant papers. After assessment of relevance and quality, only 3 articles met the minimum criteria. Heterogeneity and flaws in quality meant it was not possible to make specific inferences for practice from the studies. Large-scale rigorous research, in the form of randomized controlled trials, is urgently needed to identify whether activity-based group work is effective.


A survey of group work in occupational therapy in mental health identified that all the occupational therapists who responded used groups in their practice (Duncombe & Howe, 1985). This group work is often activity based (Lloyd, King, & Bassett, 2002), with activity defined as “a series of linked episodes of task performance by an individual which takes place on a specific occasion during a finite period for a particular reason” (Creek, 2003, p. 49). According to Finlay (2004), “activity groups aim to develop skills and/or encourage social interaction. The term is used in contrast to ‘support group’ which emphasizes communication and psychotherapy elements” (p. xiii). A wide range of benefits are attributed to participation in activity-based groups (see Arnold, 2009; Breines, 1995; Moll & Valiant Cook, 1996; Rebeiro & Cook, 1999), which involve group members’ actively doing rather than being more passive participants or using verbal skills alone. Occupational therapists believe that activity groups are more effective than verbal groups for developing interpersonal skills (Denton, 1987), because their assumption is that “human beings define their lives, cultures, values, and worth through activities” (Breines, 1995, p. 3) and that doing and communicating can be powerful therapeutic tools (Remocker & Storch, 1992).

Finlay (1993, p. 205) reported that “a lot of research on groups has been done, but findings are often contradictory and the diversity of groups and outcomes studied can be confusing. Moreover, little of the research relates directly to occupational therapy practice.” Much of the information published is descriptive in nature, and studies of the efficacy of activity-based group work are often based on self-reports from service users or caregivers (Eklund, 1997; Falk-Kessler, Momich, & Perel, 1991; Webster & Schwartzberg, 1992), which, although essential, remain subjective in nature. Use of activity-based group work also clearly varies widely among occupational therapists (Bryant, 1995;
Occupational therapists need to be able to demonstrate that their interventions are clinically effective so that the finite resources available can be used to deliver the best possible outcomes for the population served (World Health Organization, 2004). We therefore posed the following question: Is activity-based group work effective in helping people with severe and enduring mental illness in community settings improve their functional ability, reduce their mental health symptoms, or both?

Method

To develop an understanding of the existing literature on activity-based group work, a systematic review was indicated. A preliminary search of existing and ongoing systematic reviews identified none of direct relevance to the research question.

Search Strategy

We used a range of sources to search comprehensively and so reduce bias. The electronic sources searched (including the Internet) are summarized in Table 1. This search was conducted by one reviewer, and eligibility screening took place by examining titles, abstracts, and key words on the basis of the inclusion and exclusion criteria (Table 2). Abstracts were scrutinized for relevance to the study. Any abstract identified as relevant, including those abstracts that did not provide sufficient information, was retrieved in full-text format to ensure that no study was excluded because of poor bibliographic or citation detail. Relevant journals were hand searched by both reviewers (Table 3), and a citation search was conducted on all full-text journal articles and key textbooks (i.e., Breines, 1995, 2004; Cole, 2005; Creek, 1996; Cronin-Mosey 1973; Denton, 1987; Fidler & Velde, 1999; Finlay, 1993, 2004; Hagedorn, 2000; Lamport, Coffey, & Hersch 2001; Long & Cronin-Davis, 2006; Molineux, 2004; Remocker & Storch, 1992; Wilson, 1992, 1996). The full-text documents retrieved were independently reviewed by both of us using the inclusion and exclusion criteria (see Table 2). An additional independent reviewer was available to resolve any disagreements.

Inclusion and Exclusion Criteria

We formulated an inclusion and exclusion checklist based on the study design, characteristics of the population being researched, the intervention taking place, and the outcome measures used (see Table 2); for the purpose of the review, we used the definitions in Table 4.

For questions of effectiveness, the preferred study design is the randomized controlled trial (RCT; Hill & Spittlehouse, 2003). However, Sackett, Rosenberg, Gray, Haynes, and Richardson (1996) have advised that in cases in which no randomized trial has been carried out, it is best to follow the trail to the next best external evidence and continue from there. We recognized that selection bias is increased by including a wider range of study designs, but we included the following designs: RCTs, quasi-randomized studies, cohort studies, case–control studies, and controlled before-and-after studies. We did not include other types of study, such as observational studies with no control group, qualitative studies, and expert opinion, because they do not address questions of effectiveness (Roberts & DiCenso, 1999). To keep the study manageable, we focused on enduring mental illness, including the major mental illnesses of schizophrenia and psychosis, bipolar disorder, depression, and anxiety, but not organic mental disorder. We set no inclusion or exclusion criteria based on duration of illness, ethnicity, gender, or upper age limit so as to capture data for the diverse range of demographic variables this population has and to increase the external validity of the results found.

Studies were excluded if they involved groups that were verbally based (i.e., “talking about”) rather than activity based (e.g., social skills training, psychoeducational intervention, cognitive–behavioral therapy, compliance therapy, and counseling; see Table 4). Studies of specific activities (e.g., football, weaving, drumming—studies focusing on the effectiveness of one activity type) were excluded because our study’s focus was the benefit of activity-based intervention in general. We included studies using measures of functional ability, mental health symptoms, or both. Table 2 lists some of the outcome measures that were appropriate for inclusion. This list was not exhaustive because of the number of measures available within the clinical area of mental health; however, the validity and reliability of outcome measures was considered during the quality assessment. Studies found in languages other than English were excluded because of the lack of resources for translation.

Quality Assessment and Data Extraction

Once we identified full-text studies for inclusion, we conducted quality assessment using the Critical Appraisal Skills Program (CASP; 2006a, 2006b, 2006c) checklist appropriate to each study design, examining components of the methodology’s quality (e.g., internal and external validity, rather than summary scores). We each completed quality assessment; unmasked independent quality assessment by more than one reviewer is generally accepted and should be sufficient in making judgments about study quality (National Health Service Centre for Reviews
Table 1. Summary of the Electronic Sources Searched

<table>
<thead>
<tr>
<th>Database and Time Period</th>
<th>Search Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CINAHL (1982 to March 2009, Wk 4; hosted via Ovid)</td>
<td>• mental health AND group*, AND community, NOT primary care, AND activit*</td>
</tr>
<tr>
<td>Medline (1966 to March 2009, Wk 4)</td>
<td>• mental illness, AND group*, AND community, NOT primary care, AND activit*</td>
</tr>
<tr>
<td>Embase (1980 to March 2009, Wk 4)</td>
<td>• schizophrenia AND (group* OR activit*), AND community, NOT (cancer OR heart OR dementia OR stroke), NOT primary care, NOT (pharm* OR HIV), AND (severe OR enduring)</td>
</tr>
<tr>
<td>PsycINFO (1872 to March 2009, Wk 4)</td>
<td>• bipolar AND (group* OR activit*), AND community</td>
</tr>
<tr>
<td>AMED (1985 to March 2009)</td>
<td>• manic depression AND (group* OR activit*), AND community</td>
</tr>
<tr>
<td>HMIC (March 2009)</td>
<td>• depression AND (group* OR activit*), AND community, NOT (heart OR cancer OR dementia OR stroke), NOT (child* OR adolescent), NOT primary care, NOT (pharm* OR HIV), AND (severe OR enduring)</td>
</tr>
<tr>
<td>BIOSIS Previews (2004 AND 2005 Wk 1–13; hosted via Ovid; <a href="http://www.ovid.com/site/index.jsp">www.ovid.com/site/index.jsp</a>)</td>
<td>• psychosis AND (group* OR activit*)</td>
</tr>
<tr>
<td>Cochrane Library Health Technology Assessment Database (1800–March 2009; via <a href="http://www.nelh.nhs.uk/">www.nelh.nhs.uk/</a>)</td>
<td>• anxiety AND (group* OR activit*), AND community, NOT (heart OR cancer OR dementia OR stroke), NOT primary care, NOT (pharm* OR HIV), NOT (child* OR adolescent), AND (severe or enduring)</td>
</tr>
<tr>
<td>National Research Register (2000; archive); archived September 2007 (<a href="http://www.nelh.nhs.uk/Pages/NRRArchive.aspx">www.nelh.nhs.uk/Pages/NRRArchive.aspx</a>)</td>
<td>• mental health AND group(*)</td>
</tr>
<tr>
<td>Cochrane Database of Systematic Reviews (1800–March 2009; <a href="http://www.nelh.nhs.uk/">www.nelh.nhs.uk/</a>)</td>
<td>• mental illness AND group(*)</td>
</tr>
<tr>
<td>Cochrane Central Register of Controlled Trials (1800–March 2009; <a href="http://www.nelh.nhs.uk/">www.nelh.nhs.uk/</a>)</td>
<td>• schizophrenia AND group(<em>)/(group</em> AND activit*)</td>
</tr>
<tr>
<td>Cochrane Database of Abstracts of Reviews of Effectiveness (1800–March 2009; <a href="http://www.nelh.nhs.uk/">www.nelh.nhs.uk/</a>)</td>
<td>• psychosis AND group(<em>)/(group</em> AND activit*)</td>
</tr>
<tr>
<td>ISI Web of Science 5.0 (powered by ISI Web of Knowledge; <a href="http://wok.mimas.ac.uk/">http://wok.mimas.ac.uk/</a>)</td>
<td>• bipolar AND group(*)</td>
</tr>
<tr>
<td>OTseeker (March 2009; via <a href="http://www.otseeker.com/">www.otseeker.com/</a>)</td>
<td>• manic depression AND group(*)</td>
</tr>
<tr>
<td>Catalogue of the College of Occupational Therapists Library (which includes the Thesis Collection; 1979–March 2009)</td>
<td>• depression AND group(<em>)/(group</em> AND activit*)</td>
</tr>
<tr>
<td></td>
<td>• anxiety AND group(<em>)/(group</em> AND activit*)</td>
</tr>
</tbody>
</table>

Note: Asterisks were used as a wild card symbol for truncation.

and Dissemination [NHS CRD], 2001). An independent reviewer was available to resolve any disagreements; after quality assessment, we each conducted data extraction using the data extraction form for an effectiveness review (NHS CRD, 2001). The extracted data are summarized in Supplemental Table 1 (available online at www.ajot.ajotpress.net; navigate to this article, and click on “supplemental materials”).

Data Analysis

We collated and summarized the findings of the included studies to allow us to examine their similarities and differences to assess whether they were homogeneous and able to be synthesized. We could then consider calculating an average estimate of effectiveness, either overall or for subgroups of the data, and whether meta-analysis combining outcomes could be performed. Any problems highlighted because of a lack of important information were identified at this stage, and the reviewer completed an additional check to ensure that any multiple publications of the same data were included in the analysis only once to avoid bias in estimation of effectiveness. If quantitative synthesis was not possible, we planned to use a narrative synthesis to explore variations and make a judgment in relation to the findings’ value and applicability (NHS CRD, 2001).

Results

The search identified 136 papers as “potentially relevant” to this study. We excluded 108 immediately because they did not meet inclusion criteria; for example, interventions were not purely activity based (Norton & Smith, 2005),
Table 2. Study Inclusion and Exclusion Checklist

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Yes/No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inclusion</strong></td>
<td></td>
</tr>
<tr>
<td>Research population (all criteria need to be met for inclusion)</td>
<td></td>
</tr>
<tr>
<td>Have a primary diagnosis of severe-enduring mental illness?</td>
<td></td>
</tr>
<tr>
<td>Receiving services and living in the community?</td>
<td></td>
</tr>
<tr>
<td>Age ≥18?</td>
<td></td>
</tr>
<tr>
<td>Study design (one of these criteria to be met for inclusion)</td>
<td></td>
</tr>
<tr>
<td>Systematic review?</td>
<td></td>
</tr>
<tr>
<td>Randomized controlled trial?</td>
<td></td>
</tr>
<tr>
<td>Quasi-randomized study?</td>
<td></td>
</tr>
<tr>
<td>Controlled clinical trial?</td>
<td></td>
</tr>
<tr>
<td>Controlled before and after study?</td>
<td></td>
</tr>
<tr>
<td>Intervention (all criteria to be met for inclusion)</td>
<td></td>
</tr>
<tr>
<td>Does intervention consist of “doing” activities?</td>
<td></td>
</tr>
<tr>
<td>Is intervention completed in a group?</td>
<td></td>
</tr>
<tr>
<td>Outcome (at least one of the following to be used for inclusion)</td>
<td></td>
</tr>
<tr>
<td>Comprehensive Mental Health Assessment</td>
<td></td>
</tr>
<tr>
<td>Brief Psychiatric Rating Scale</td>
<td></td>
</tr>
<tr>
<td>Beck Depression Inventory (or similar for anxiety—Beck Anxiety Inventory)</td>
<td></td>
</tr>
<tr>
<td>Allen Cognitive Level Screen</td>
<td></td>
</tr>
<tr>
<td>Comprehensive Occupational Therapy Evaluation</td>
<td></td>
</tr>
<tr>
<td>Global Assessment Scale</td>
<td></td>
</tr>
<tr>
<td>Social Avoidance and Distress Scale</td>
<td></td>
</tr>
<tr>
<td>Are other outcome measures used?</td>
<td></td>
</tr>
<tr>
<td><strong>Exclusion</strong></td>
<td></td>
</tr>
<tr>
<td>Research population (exclude if any of these criteria are met)</td>
<td></td>
</tr>
<tr>
<td>Have a diagnosis of organic mental illness (e.g., Alzheimer’s disease, dementia)?</td>
<td></td>
</tr>
<tr>
<td>Have a dual diagnosis (i.e., mental illness and substance abuse issues)?</td>
<td></td>
</tr>
<tr>
<td>Have mild to moderate mental illness?</td>
<td></td>
</tr>
<tr>
<td>Receive mental health intervention via primary care services?</td>
<td></td>
</tr>
<tr>
<td>Study design (exclude if any of these criteria are met)</td>
<td></td>
</tr>
<tr>
<td>Qualitative study?</td>
<td></td>
</tr>
<tr>
<td>Case study?</td>
<td></td>
</tr>
<tr>
<td>Descriptive study?</td>
<td></td>
</tr>
<tr>
<td>Expert opinion?</td>
<td></td>
</tr>
<tr>
<td>Intervention (exclude if any of these criteria are met)</td>
<td></td>
</tr>
<tr>
<td>Is the intervention related to one specific activity?</td>
<td></td>
</tr>
<tr>
<td>Is the intervention counseling?</td>
<td></td>
</tr>
<tr>
<td>Is the intervention cognitive-behavioral?</td>
<td></td>
</tr>
<tr>
<td>Is the intervention social skills training?</td>
<td></td>
</tr>
<tr>
<td>Is the intervention psychoeducational?</td>
<td></td>
</tr>
<tr>
<td>Does the intervention involve “talking about” rather than “doing”?</td>
<td></td>
</tr>
<tr>
<td>Is the intervention compliance therapy?</td>
<td></td>
</tr>
<tr>
<td>Outcome (exclude if any of these criteria are met)</td>
<td></td>
</tr>
<tr>
<td>There are no outcome measures</td>
<td></td>
</tr>
</tbody>
</table>

study populations did not focus on severe and enduring mental illness (Cohen et al., 2006), or research design was not relevant (Nose, 2005; full details are available from Alison Bullock). We each read the full text of 27 of the remaining studies and applied the inclusion–exclusion checklist, although 1 study remains unaccessed. After this process, 24 studies were excluded. Examples of reasons for exclusion included that the study design did not meet inclusion criteria (Ardath & McDermott, 1988), the intervention related to a discrete activity (Bickes, DeLoache, Dicer, & Miller, 2001), and the intervention was verbal rather than activity based (Daniels & Roll, 1998).

We agreed on all exclusions without recourse to the independent reviewer. Three studies were identified as eligible for review (DeCarlo & Mann, 1985; Klyczek & Mann, 1986; Schindler, 1999) and were appraised for quality using the appropriate CASP checklist for each study design (CASP, 2006a, 2006b, 2006c), and data were extracted. Both of us completed this process. Some homogeneity between the studies was found. All study participants were attendees at a mental health day treatment center or day hospital, all involved an adult population, all used verbal therapy and discussion as a comparison group, and all were conducted in the U.S. health care system.

The high level of heterogeneity meant that it was not possible to use meta-analysis to calculate an average estimate of effectiveness either overall or for subgroups of the data; thus, a quantitative analysis was inappropriate, and a narrative review was required. Examples of heterogeneity included variable lengths of intervention (from 1 hr/wk to 4.2 hr/wk), use of different outcome measures, and the application of different inclusion and exclusion criteria. All three studies had flaws in internal validity, such as failing to detail randomization (DeCarlo & Mann, 1985), lack of baseline characteristics to ascertain whether groups were well balanced (DeCarlo & Mann, 1985; Schindler, 1999), using tools without demonstrated reliability and validity (Klyczek & Mann, 1986), identifying some confounding factors but failing to collect data in relation to those factors (DeCarlo & Mann, 1985; Schindler, 1999), and not describing blinding (DeCarlo & Mann, 1985).

That cost-effectiveness was not described is unsurprising when one considers the age of the studies included; cost-effectiveness has been a more recent concern.

**Narrative Review**

All the studies showed that activity-based group work was more effective than verbally based group work, but for different reasons—that is, improving self-perceptions of social interaction skills, improving social behavior, providing greater symptom reduction, and increasing levels of community functioning. The results of the studies with social interaction outcomes were statistically significant (both ps < .05; DeCarlo & Mann, 1985; Schindler,
for activity-based groups compared with verbally based groups; however, no inferential statistical testing was completed in the study that indicated greater symptom reduction and increased levels of community functioning for clients receiving more activity therapy (Klyczek & Mann, 1986). Schindler (1999) found a statistically significant difference in social functioning outcomes between the activity-based group and the control group; however, DeCarlo and Mann (1985) discovered no significant difference when comparing an activity-based group with a control group. Both studies had small sample sizes (19 and 25, respectively). Klyczek and Mann (1986) demonstrated equivalent community tenure (i.e., the number of days living in the community) between participants receiving twice as much activity therapy as verbal therapy and those receiving twice as much verbal therapy as activity therapy, but a greater relapse rate occurred for those in the primarily activity-based group (although not to the point of hospital admission).

Generalizing from any of these studies is not possible because of lack of rigor, level of heterogeneity, and small sample sizes, despite two of the studies’ finding statistically significant effects (DeCarlo & Mann, 1985; Schindler, 1999). Although it is encouraging that positive effects were observed, the results of the review do not provide sufficient evidence to either support or refute the effectiveness of activity-based group work in either improving functional ability or reducing mental health symptoms of community-living adults with severe and enduring mental illness.

Discussion

The findings of this systematic review are problematic for occupational therapists because of the extensive use of activity-based groups in occupational therapy practice. Consequently, our discussion emphasizes that the lack of evidence does not mean that no effect or directions for future research exist. These points are discussed in turn.

Lack of Evidence Is Not Evidence of No Effect

Although no evidence exists to either support or refute the effectiveness of activity-based group work in either improving functional ability or reducing mental health symptoms of community-living adults with severe and enduring mental illness, this lack of evidence does not mean that activity-based group work is ineffective. Rather, it means that we do not know one way or the other whether activity-based groups are effective, a subtle
but important distinction. This situation also reflects the fact that rigorous evidence is limited for any non-pharmacological interventions that are effective in a population with severe and enduring mental illness; however, such evidence is building (Buchain, Vizotto, Neto, & Elks, 2003; Scott, Palmer, Paykel, Teasdale, & Hayhurst, 2003). The literature has shown that a substantial amount of consensus exists with regard to the benefit of activity-based group work across professionals, service users, caregivers and family, and mental health organizations (Eklund, 1997; Moll & Valiant Cook, 1996; Willcock, 2006). However, “because patients so often get better or worse on their own, no matter what we do, clinical experience is a poor judge of what does and does not work” (Doust & Del Mar, 2004, p. 474). Robust research evidence is needed to assess whether the consensus on the benefit of activity-based group work is accurate.

**Directions for Future Research**

Rigorous, well-powered studies need to be conducted to test ideas about the benefits and effectiveness of activity-based group work, and the studies included in this review demonstrate that it is possible to complete RCTs and quasi-experimental studies in this clinical area. Occupational therapy interventions used with other client groups (e.g., older people with and without dementia) have also been rigorously researched and shown to be effective (Clarke et al., 1997; Graff et al., 2006), findings that bode well for research about activity-based group work with people who have severe and enduring mental illness. The literature has highlighted the greater need for occupational therapy research in the mental health field (Bannigan, Boniface, Doherty, Porter-Armstrong, & Scudds, 2008; Craik, Austin, Chacksfield, Richards, & Schell, 1998; Rebeiro 1998). Attention also needs to be given to the methodological considerations highlighted by this study when designing research to assess the effectiveness of activity-based interventions.

**Methodological Considerations for Effectiveness Research**

Since the studies included in the review were published, the Consolidated Standards of Reporting Trials (CONSORT) statement (Altman et al., 2001; Begg et al., 1996) has been developed. This statement aims to improve the quality of reporting of RCTs in particular. Occupational therapists involved in designing and reporting effectiveness studies should now consider the CONSORT statement when designing and publishing RCTs. Although the studies included in this review were completed before the CONSORT statement and other guidelines, we can learn from them how to design better studies to assess the effectiveness of activity-based group work in the future. The key issues are the lack of rigor, the high level of heterogeneity across the studies, and the lack of an intervention schedule. If those issues are addressed, the quality of future studies will be improved and will be more likely to provide evidence to support practice.

**Need for Increased Rigor in Primary Studies**

Attrition rates affect whether a sample remains representative, and the attrition rate was problematic in two of the studies found. Klyczek and Mann (1986) gave no information on attrition. Schindler (1999) gave raw figures for attrition, which was highest in the activity-based group (50%) and described as being the result of discharge. Exploring why attrition may have been so high in the activity group would have been particularly useful. Other issues include a lack of 6-mo follow-up, lack of adequate control and collection of data regarding confounding factors, and a lack of confidence intervals to aid inference (see Supplemental Table 1). These types of methodological issues should now be considered an essential part of rigorous research reporting in occupational therapy.

Occupational therapists also need to be aware of new developments in research. DeCarlo and Mann (1985), Klyczek and Mann (1986), and Schindler (1999) could not have been expected to consider service user involvement; however, it is now an essential requirement for studies and should be part of any future research designs (Department of Health (UK), 2006; Diamond, Parkin, & Morris 2003; Telford & Faulkner, 2004).

**Reducing Heterogeneity Between Studies**

We have already noted that the many differences among the primary studies make it impossible to make comparisons or pool the findings of studies using meta-analysis in secondary research, which implies that more agreement is needed among occupational therapists about population characteristics, the aspects of intervention, and the outcome measures used. In relation to population characteristics, two points need to be considered: First, which groups of patients constitute people with severe and enduring mental illness? No widely accepted agreement exists about this term, and different groups of patients may have been included, making it difficult to compare studies and to make judgments about whether the sample is representative. Consensus is necessary with regard to the terminology used to describe populations.
The second issue from a methodological perspective is that two studies in this review did not ascertain whether groups were similar at baseline (DeCarlo & Mann, 1985; Schindler, 1999). Having comparison groups that are similar at baseline is a precondition for effectiveness studies; if groups are not balanced for relevant baseline characteristics, the differences in outcomes cannot be attributed with confidence to the effects of the intervention of interest (NHS CRD 2001). Similarity of comparison groups at baseline should be part of effectiveness studies as a matter of course.

During the review, it also became apparent that more studies are available related to people with schizophrenia than to people with other diagnoses, such as severe and enduring anxiety (pharmacological interventions aside). McDermott (1988, p. 87) stated that “according to reports in the literature, chronic schizophrenic individuals differ from non-psychotic individuals in their responses to various treatment approaches,” and DeCarlo and Mann (1985, p. 22) reported that “there are qualitative differences in communication between normal and schizophrenic patients.” Such differences are one explanation for these patients being singled out for specific study and gives support to this user group’s being researched separately from those with other severe and enduring mental health diagnoses or indicates that this population should be explored as a confounding variable in mixed studies. One of the studies included in this review used only participants with schizophrenia (another factor affecting whether results can be generalized), and the other two involved participants with a mixture of diagnoses, although diagnosis was not examined as a confounding variable.

A wide range of outcome measures is used within mental health, and approximately 50 different outcome measures were used across the included and excluded studies. Some measures were formulated or adapted specifically for the study, and only a few were used in more than one study, for example, Goal Attainment Scaling (Eklund, 1999; Hayes, Halford, & Varghese, 1991; Schwartzberg, Howe, & McDermott, 1983) and the Positive and Negative Syndrome Scale (Dobson, McDougall, Busheikin, & Aldous, 1995). Analysis of the outcome measures used highlights a need for sourcing and developing measures that are effective, valid, and reliable in describing and measuring occupational therapy outcomes in mental health (Craik et al., 1998; Department of Health [UK], 1999; Rebeiro, 1998) because of the use of tools that lack reliability and validity has implications for the applicability of these studies’ results.

The measures used also clearly show a lack of emphasis on the wider benefits of activity-based group work as detailed in the background literature. For example, Webster and Schwartzberg’s (1992) study of patients’ perceptions of curative factors in groups found similar results for activity groups as for psychotherapy groups in relation to Yalom’s 11 curative factors of group therapy (additional factors of accomplishment, relaxing or keeping one’s mind off problems, creativity, self-esteem, enjoyment or fun, and increased skills and concentration were also identified). Likewise, Rebeiro and Cook (1999, p. 178) described outcomes related to affirmation, acceptance, belonging, normalization with group, self-worth, confirmation, alternate focus for thoughts, accomplishment, hope for the future, confidence, and actualization (redefining of self based on experience) as being important to service users. Observations such as these were not reflected in the outcomes identified and analyzed in this review. This phenomenon has previously been noted by Thompson and Blair (1998, p. 58), who reported that “more research is clearly needed as these papers only highlight a few benefits of activity groups.” Outcome measures for activity-based group work need to consider outcomes at all levels, from the ability to use skills to complete specific tasks to the ability to participate more fully in life roles and occupations.

### Importance of the Intervention Schedule

The intervention schedule is fundamental and should not be ignored; it should be reported in sufficient detail to be replicable in any effectiveness study.

When discussing activity-based group work as an intervention, one acknowledges difficulties with researching interventions in group settings. Clear, replicable intervention schedules are vital to reduce the effect of confounding factors. Finlay (1993) reported that by definition group processes are multiple and complex, and this creates difficulties for researchers who are trying to control variables and disentangle effects. Further, no two groups are identical and this makes it hard to compare groups or even replicate studies. (p. 205)

Confounding factors resulting from this complexity (e.g., relationships between participants and between participants and facilitator, changes to group structure and number, significant events in participants’ lives including changes in medication and cointerventions) were briefly acknowledged across the included studies but were not further explored.

Another key issue for occupational therapists is the importance of occupation-focused interventions, which lead to a wide range of benefits. Two of the included studies focused on activities to improve social interaction outcomes.
only (DeCarlo & Mann, 1985; Schindler, 1999), and 10 of the excluded studies also focused mainly on this area (Ardath & McDermott, 1988; Daniels & Roll, 1998; Mumford, 1974). Describing and monitoring the interventions received by control groups—that is, social milieu therapy and free-choice activity—are also important. For example, the control group in DeCarlo and Mann (1985) may have participated in similar amounts of group activity-based therapy as the activity group.

The concept of an activity’s meaningfulness also introduces a confounding variable and requires attention because an activity’s meaningfulness is an important feature in the successful use of activity to improve health (Breines, 1995; Denton, 1987; Moll & Valiant Cook, 1996). It is apparent in this review that one of the activity groups (Schindler, 1999) had a clear purpose related to the participants’ daily lives, whereas another (DeCarlo & Mann, 1985) had a purpose related to improving skills that can then theoretically be transferred to daily life. Schindler (1999) found a statistically significant difference between the activity-based group and both the control group and the verbally based group, whereas DeCarlo and Mann’s (1985) results showed a statistically significant difference only between the activity-based group and the verbally based group. (Greater improvements were found in the activity-based group than in the control group but not at a statistically significant level.) A larger sample may have clarified this result, but results may also relate to the meaningfulness attributed to the activities completed by participants, and meaningfulness should be an integral component of any future research in this area.

Limitations of the Study

Although we implemented a wide-ranging search strategy, our search of the gray literature was limited. The only unpublished data accessed were from the College of Occupational Therapists Theses Collection available via The College of Occupational Therapists library catalog, which we both had access to because we are members of the British Association of Occupational Therapists. This limitation may have introduced publication bias because only studies indicating an association are likely to be published (Bowling, 2002). The search of unpublished data that we conducted identified only one additional study for consideration, suggesting that no large volume of unpublished data has been missed and so is unlikely to have biased this study. The guidelines published in relation to individual conditions by the National Institute for Health and Clinical Excellence (NICE) in the United Kingdom have not identified studies related to activity-based group work, a situation that also supports the view that the limited search of gray literature did not adversely affect this study (NICE, 2004a, 2004b, 2006, 2009).

Another potential source of bias was the exclusion of non–English language studies; one study published in a language other than English was identified but excluded because we lacked the financial resources to fund translation, and we do not know what this study would have contributed to our findings. We acknowledge that the use of one researcher to screen titles and abstracts may have introduced bias. No attempt was made to contact the studies’ authors for missing data because it has been >20 yr since completion of two of the studies (DeCarlo & Mann, 1985; Klyczek & Mann, 1986); however, because conducting a meta-analysis was not possible, the missing data did not have a material effect on the study’s findings.

Conclusion

This study has highlighted the lack of rigorous scientific research to support the practice of activity-based group work with this user group. Although three studies met the minimum criteria of the review and showed that activity-based group work was more effective than verbally based group work, the level of heterogeneity meant that quantitative synthesis was prohibited. The narrative analysis concluded that, given the review results, generalizing to external populations or making definite clinical inferences was inappropriate. The most that can be concluded is that activity-based group work may well be an effective intervention of similar or improved effect as verbally based groups with similar aims. Because activity-based groups are widely used in practice, our recommendations are that this use should remain unchanged on the basis of review results, although the need for further large-scale rigorous research is identified as a matter of urgency. If this research is not conducted, these interventions will not be included in evidence-based practice guidelines and so are unlikely to be commissioned in future. ▲

Acknowledgment

This study was originally completed in partial submission of an MSc in evidence-based practice at the University of Teesside, UK. We have since updated it.

References


+ Indicates studies that were systematically reviewed for this article.