

OUTCOME EVALUATION OF A BRIEF SHARED LEARNING PROGRAMME IN COGNITIVE BEHAVIOURAL THERAPY

Pamela J. Myles

Newcastle, North Tyneside and Northumberland Mental Health NHS Trust, UK

Derek L. Milne

University of Newcastle Upon Tyne, UK

Abstract. The purpose of this study was to evaluate a new brief training programme in Cognitive Behavioural Therapy (CBT) within a waiting list control group design. A consecutive sample of 90 mental health professionals were allocated routinely by their managers to be trained in CBT. Over 12 week training periods, these 90 professionals received a total of 48 hours of shared learning, covering core CBT theory and techniques for depression and anxiety. An experiential workshop format was utilized throughout and a comprehensive outcome evaluation was conducted, using multiple measures and methods. The findings indicated that the training was effective in terms of the favourable reactions of the participants, the knowledge and skills they gained, and the generalization of this learning to their work environment. It is concluded that brief shared training in CBT can be successful, in line with current NHS priorities.

Keywords: Cognitive behavioural therapy, shared learning, evaluation.

Introduction

Mental health problems are common, depression being particularly widespread. At any one time, one in six adults suffers from a form of mental illness such as anxiety or depression, while more severe mental illness, such as schizophrenia, affects less than 1% of the population (MACA, 1999). The National Service Framework for Mental Health (Department of Health, 1999) sets standards, to ensure that any service user with a common mental health problem has their mental health needs identified and assessed, and is then offered effective (i.e. evidence-based) treatment. This includes a national programme of training to address critical skills gaps. It follows that, for this standard to be met, we must improve the training of mental health professionals in empirically supported treatments (Roth & Fonagy, 1996; Freiheit & Overholser, 1997).

Because of its strong research base and testable interventions, Cognitive Behavioural Therapy (CBT) represents a therapeutic approach that is highly compatible with evidence-based

Reprint requests to Pamela Myles, Psychological Services Department, West Farm House, 1–2 West Farm Court, Station Road, Cramlington, Northumberland NE23 1AX, UK. E-mail: pamela.myles@nmht.nhs.uk

© 2004 British Association for Behavioural and Cognitive Psychotherapies

practice (EBP). The implementation of EBP is likely to increase the demand for CBT training, particularly for brief, shared learning experiences that are themselves evidence-based.

Methodological issues

In general terms, the literature on the effectiveness of training in psychotherapy (regardless of therapeutic orientation) is inconclusive (Roth & Fonagy, 1996; Bootzin & Ruggill, 1988; Strupp, Butler, & Rosser, 1988; Wright, Horlick, Bouchard, Mathieu, & Zeichner, 1977). This is partly due to methodological weaknesses in the extant research, and partly to the considerable difficulties in procuring the generalization of training to the workplace (Milne, Keegan, Westerman, & Dudley, 2000). To be specific, these authors have criticized research for a number of methodological weaknesses, including small sample sizes, problems with external validity (e.g. analogue measures) and problems with internal validity (e.g. research designs have typically taken the form of simple pre-post designs without any controls), resulting in only limited conclusions being possible. In addition, measurement typically does not allow for an analysis of whether transfer of training occurs at all, or is temporary (Oliver & Fleming, 1997).

Jahr (1998) has suggested that the overall effectiveness of staff training programmes should be evident ultimately through the trainees' ability to transfer the skills learned across settings. He argued that researchers should measure this transfer of training, as "there is no guarantee that skills acquired in a workshop setting will be applied correctly, or at all, in the daily work setting" (p. 78).

A later critical review of skills training studies by Alberts and Edelstein (1990) found that little progress had been made in the methodological sophistication of studies and that deficiencies existed regarding target skill selection, measurement, and training procedures. They concluded that "investigators must aspire to higher standards of basic research design and measurement considerations if the knowledge base for our training technology is to advance" (p. 509). In particular, they noted the need for multiple measures. Additionally, Oliver and Fleming (1997) have made several recommendations for more rigorous research on the transfer of training, including the use of repeated measures.

Generalization issues

Turning to the issue of the generalization (or "transfer") of training, it has been found repeatedly that the clinical environment to which the trainee returns has an important impact on the transfer of training. When trainees return to work there are a number of constraints they may encounter that will serve to reduce the transfer of what has been learned. From a small but empirical study, Milne, Gorenski, Westerman, Leck and Keegan (2000) identified nine factors affecting the generalization of training, five of which benefited and four of which hindered generalization. The beneficial factors included support from colleagues, benefits to clients, knowledge gained, supervision and regular meetings. Barriers included the lack of available time, lack of staff, lack of resources and not enough staff trained in the procedures.

Training in CBT

Research into the effectiveness of CBT training is also relatively limited. Nevertheless, in the available literature there is evidence of the effectiveness of training in improving CBT

competence amongst professionals. For example, an intensive training programme of 12–18 months (with supervision continuing to 24 months) was evaluated by Shaw and Wilson-Smith (1988). Therapy sessions of 10 trainees were monitored by independent raters, using the Cognitive Therapy Scale (CTS) (Young & Beck, 1980, 1988). Over time, significant training effects were produced on the overall CTS scores, on the conceptualization sub-scale, and on the structure of CBT score (e.g. agenda setting, pacing and homework).

Subsequently, Williams, Moorey and Cobb (1991) evaluated the impact of a year-long CBT course on competence levels pre and post-training. Audiotaped clinical sessions were rated blind to condition, using a shortened form of the CTS (Young & Beck, 1980, 1988). Eight of the 11 trainees showed improvements in CTS scores following training. Improvement in competence (as rated by CTS total scores) just failed to reach significance. However, the global competency rating showed significant mean improvement from pre to post-training. A major flaw in the study is that the raters were also the trainers, and so were not independent.

Freiheit and Overholser (1997) examined the impact of therapeutic orientation on knowledge, attitudes and use of CBT skills over time, following a year-long course in CBT provided to 40 clinical psychology graduate students. Trainees gained a significant amount of knowledge from pre to post-testing. Attitudes about cognitive ideology significantly increased, while there was no change in behavioural ideology. There was also a significant reduction in negative appraisal of CBT, and a significant increase in the use of cognitive and behavioural techniques, regardless of original therapeutic orientation. Unfortunately, the study was limited in having used only clinical graduate students, who may have been more open to different orientations. A second criticism is the use of only one measure, which was also unvalidated.

Milne, Baker, Blackburn, James and Reichelt (1999) evaluated the impact of a year-long training programme in CBT on the competence of 20 mental health professionals using a revised version of the CTS (CTS-R). Levels of competence were then related to the clinical outcomes obtained by the trainees' patients. They reported that training produced a highly significant increase in overall therapist competence: 9 of the 14 CTS-R subscales indicated significant improvements in therapist competence. They also found improvements in the patients' coping strategies (i.e. more frequent use of positive appraisal and problem solving and less frequent use of cognitive avoidance and emotional discharge following therapy). However, only the reduction in emotional discharge produced statistically significant results. Inclusion of assessment of tapes prior to any training would have strengthened the study further.

King, Davidson, Taylor, Haines, Sharpe and Turner (2002) describe a randomized controlled trial of a CBT training package using 84 GPs. They found no discernible impact on patient outcomes and no major difference between intervention and control groups on attitudes towards its treatment.

In summary, there is a lack of rigorous research on training in CBT. Of those studies discussed, most used simple pre-post designs, resulting in limited conclusions being possible. Also, studies rarely measure the generalization of training, whether in terms of maintenance or the benefits to patients. A further weakness has been the reliance on only one or two learning outcomes. The present training programme attempts to address these methodological weaknesses. The evaluation takes the form of a waiting list control group design, i.e. a longitudinal study, with participants serving as their own controls. The key evaluation components were the comprehensive measurement of the learning impacts of the course. In addition, multiple measures were administered over four time-points, drawing on multiple

methods. It was predicted that this brief training programme would result in significant learning effects and that there would be high levels of learner satisfaction. However, because there had been no systematic programming of generalization, we expected this to be poor.

Method

The cognitive-behavioural therapy training programme was based and accredited at the University of Northumbria at Newcastle, and run over a 12-week period. Each of the 12 once-weekly sessions took the form of a 4-hour workshop. The lead trainer (the first author) submitted a videotape of a typical teaching session in order that the quality of teaching could be examined by an independent assessor (the second author) using an observation manual (Milne, James, Keegan, & Chalk, 2001). This assessment found that an appropriately large proportion of the training analysed consisted of role-play and discussion. A second trainer (a lecturer from Northumbria University) assisted with the training at times and guest presenters from Newcastle, North Tyneside and Northumberland Mental Health NHS Trust with expertise in CBT contributed to a number of sessions.

The course aimed to develop a range of basic skills to enable the participants to work within a well-established cognitive-behavioural framework and to apply basic techniques to patients with Axis I disorders (e.g. mild to moderate depression and commonly encountered anxiety disorders, such as Generalized Anxiety Disorder and Panic Disorder).

Participants

Six cohorts of 15, 17, 15, 15, 16 and 12 (total $n = 90$) participants were included. The groups consisted of registered mental nurses ($n = 76$), occupational therapists ($n = 5$), counsellors ($n = 2$), social workers ($n = 2$), psychology assistants ($n = 2$), psychiatrists ($n = 2$), and a registered mental handicap nurse ($n = 1$). There were 69 females and 21 males, with an average age of 36.84 years. Participants came from Trusts throughout the northern region. There were 12 drop-outs from the training due to a variety of reasons (e.g. job changes/demands; illness/accident; bereavement). Of the remaining 78 participants, 36 agreed to attend the university to complete pre-training assessment (first baseline) and 55 completed measures at 3-month follow-up.

Procedure

Entry to the teaching programme was by open competition amongst mental health staff in the northern region provided that they possessed a professional qualification and had direct clinical contact with patients. It was not necessary for them to have specialist knowledge of CBT or of working within this framework. No interviews took place; if the individual had the appropriate qualifications, they were offered a place on the course on a first come, first serve basis.

Assessments of learning by the participants took place at four time-points. There was a first baseline (3-months pre-training), a second baseline (first training session), a post-training assessment (end of the final training session) and a 3-month follow-up. The first and second baselines enabled us to treat the group as their own controls, as in a "waiting list" condition. This strengthened the research design, as we were able to determine the effect of a number of potential threats (e.g. repeated testing and the passage of time). The assessment tasks were

Table 1. Summary of the instruments used to evaluate the CBT training course

Dimension evaluated	Instruments used	Time point
Learner satisfaction	Training Acceptability Rating Scale	3 (final session)
Knowledge	a. Multiple Choice Questionnaire	1,2,3,4 (all four time-points)
Knowledge	b. Video Assessment Task	1,2,3 (first & second baseline, final session)
Impact evaluation	a. Generalization Questionnaire	3,4 (final session and follow-up)
Impact evaluation	b. Barriers to Change Questionnaire	4 (follow-up)

conducted in a group setting and took just a few minutes to complete. The follow-up measures were sent to the participants 12 weeks post-training and on completion were returned in an enclosed stamped addressed envelope to the first author. The measures were administered by the lead trainer (first author) and coded to enable comparisons to be made at the four time points.

Materials

Measurement was made of the impact of the course, covering participants' reactions, knowledge gain, and the generalization of skills to patients. The measurement methods included self-report questionnaires and a video presented task. A number of validated questionnaires were adapted to the circumstances of the study to ensure an appropriately relevant and rigorous analysis, while other measures were created and developed for the course as necessary. These instruments are now described, in the order of administration.

Learner satisfaction. At the final session the learners' satisfaction was evaluated using the Training Acceptability Rating Scale (Part I from Davis, Rawana, & Capponi, 1989, test re-test reliability $r = 0.83$ $p < .01$, Part II adapted from Milne & Noone, 1996). Part II was altered minimally, by changing wording from "workshop" to "course". This assessment tool has 15 items in total and has been adapted for relevance to CBT training and requires the participant to rate the general acceptability, effectiveness, negative side-effects, appropriateness, consistency and social validity of the course.

Measure of knowledge. An ad hoc 13-item multiple choice questionnaire was developed to assess level of knowledge of CBT, including the therapeutic model, principles of the therapy and treatment strategies. It was administered at both baselines, post-training, and at the follow-up assessment. The respondent was expected to answer every question, each of which had four response options giving a 25% chance of guessing the correct answer (Myles & Milne, 2001). The MCQ was vetted by two experienced colleagues who agreed that the measure reflected the course content. It was taken seriously by trainees, demonstrating good content and face validity. Test-retest reliability was highly significant at $r = 0.85$, $p < .0001$ (Myles & Milne, 2001).

Measure of knowledge. The Video Assessment Task (VAT) was based on the Functional Analysis Task developed by Milne (1984), but adapted to include cognitive factors. A script was developed that included three symptoms from each of the following domains: cognitive, behavioural, physical, and affective. An actor spoke directly to camera and described her problem, which was a typical example of Panic Disorder. Participants were shown the video

clip twice and given a total of 10 minutes to answer three questions related to what they had seen. These were: identifying symptoms; ascertaining the problem; and considering up to six cognitive-behavioural techniques to use in treatment. The maximum score that could be attained was 24, the higher the score the greater the skill and knowledge demonstrated by the participant. This measure was administered at both baselines and at post-training assessment. The measure was shown to have adequate content and face validity and had highly significant inter-rater reliability (Section A $\rho = 0.97$ $p < .01$; Section B $\rho = 1$ $p < .01$; Section C $\rho = 0.94$ $p < .01$).

Impact evaluation (a). At the 3-month follow-up phase, participants were asked to complete the Barriers to Change Questionnaire (BARCQ; Corrigan, Karatanini, & Pramanan, 1992). This instrument assesses to what extent participants believe that the training they have received can be introduced as part of their routine work. A higher score indicates that factors at work hinder the trainees' application of what they have learned in their training. Corrigan et al. (1992) reported test-retest reliability of $r = 0.41$ to 0.92 $p < .006$ indicating significant test-retest reliability.

Impact evaluation (b). The extent to which the training had generalized (i.e. transferred to the work environment, and across time, responses and to clients) was measured using the Generalization Questionnaire, based on Milne, Gorenski et al. (2000). The questionnaire focused on the four main areas of Experience, Generalization, Areas of functioning, and Support and participation. This measure was administered at the final training session and at the 3-month follow-up assessment.

The "Experience" category (Section A) has five items related to the participants' practical experience in the areas addressed on the course (e.g. the CBT model). Twelve methods and instruments are listed in the "Generalization" category (Section B), which participants were asked to tick if they had used them in their clinical practice during the period of the course (e.g. monitoring diaries). Participants were then asked to circle the letter corresponding to the items they had used in the 3 months prior to the training. There are three items in the "Areas of functioning" category (Section C), which are rated on a 1–5 scale where 1 represents "No impact" and 5 represents "Extensive impact" of the training on e.g. ability to resolve clients' problems. The final category "Support and participation" (Section D) lists seven activities that the trainees may have been involved in within the work environment while attending the course (e.g. team meetings, supervision) and asked to identify the three most helpful activities.

Research design

The research utilized a waiting list, control group design. It was a longitudinal study, with participants serving as their own controls. Multiple measures of the participants' thoughts, feelings and behaviours were administered during a double baseline assessment, a post-training re-assessment and at a 3-month follow-up assessment.

Results

Training Acceptability Rating Scale

The results for the two subsections and total are: Acceptability (score 6–36) $n = 69$, median 33, (inter-quartile range 29–33.5); Effectiveness (score 0–27) $n = 69$, median 22, (inter-quartile

range 19–23); Total scores (score 6–63) $n = 69$, median 52, (inter-quartile range 49–56). These results suggest, that in the opinion of participants, the training was acceptable and effective.

Foundation Cognitive Behavioural Therapy Multiple Choice Questionnaire (FCBT-MCQ)

A Wilcoxon test was conducted to determine whether there were differences between the first baseline (3 months pre-training) and second baseline assessments. Data from 36 cases were available at double baseline. These data indicated that there was no significant difference between first baseline (median 7) and second baseline (median 7), ($z = -1.16$, $p = .25$; ns). This supports the view that the passage of time, the repetition of the measure alone, and other possible confounds did not produce higher scores.

A Friedman test was used to compare the second baseline (median 7) and post-training (median 7) MCQ data. A highly significant difference was obtained ($\chi^2(2, N = 55) = 76.18$, $p < .0001$). Follow-up pairwise comparisons were conducted using a Wilcoxon test with a Bonferroni correction, to control for Type I errors. There were significant differences between the first and final training session ($z = -6.27$, $p < .0001$) and between the first training session and 3-months post course ($z = -6.17$, $p < .0001$). Also, there was no significant difference between the final training session assessment (median 11) and the 3-month post-course follow-up assessment (median 11), ($z = -0.48$, $p = .63$). These findings indicated that participants' knowledge of CBT increased significantly with training, and was maintained at 3-month follow-up.

Video Assessment Task (VAT)

Due to changes made to the initial instrument, the following data analysis is for participants drawn from the three latter cohorts ($n = 19$) at the two baseline assessments and post-training assessment. A Friedman Test was conducted to determine whether there had been a change in knowledge using the VAT. The test indicated a significant difference on Sections A (scoring 0–24) and C (scoring 0–6), (i.e. identifying symptoms and naming up to six CBT interventions: $\chi^2 = 22.9$ and 26.15, respectively: $p < .0001$) but did not reach significance on the section (B) (scoring 0–2) concerned with problem identification ($\chi^2 = 3.94$, $p = .14$).

Follow-up pairwise comparisons were again conducted using a Wilcoxon test with a Bonferroni correction, to control for Type I errors (see Table 2 below). On Section A there was no significant difference between time-points 1 (median 4) and 2 (median 6), but significant differences were found between time-points 1 (median 4) and 3 (median 10), and time-points 2 (median 6) and 3 (median 10). Although there were no significant differences found in pairwise comparisons of Section B, there was a trend showing some improvement in identification of the problem following training. For Section C no significant differences were found between time-points 1 (median 3) and 2 (median 3), but differences between time-point 1 (median 3) and 3 (median 5) and time-point 2 (median 3) and 3 (median 5) were highly significant ($p < .0001$). These results indicate that improvement in the VAT was due to the CBT training.

Barriers to Change Questionnaire

Ratings from the Barriers To Change Questionnaire showed that the largest perceived barrier to implementing what had been learned was “Institutional constraints” ($n = 55$), median 9.5

Table 2. Results of the Wilcoxon test for the VAT ($N = 19$)

Section	Time points	Z	Probability
A (symptoms)	1 and 2	-2.43	0.02 ns
A (symptoms)	1 and 3	-3.63	0.0001
A (symptoms)	2 and 3	-3.47	0.001
B (problem)	1 and 2	0.00	1.00 ns
B (problem)	1 and 3	-1.67	0.96 ns
B (problem)	2 and 3	-1.51	0.13 ns
C (treatment)	1 and 2	-0.46	0.64 ns
C (treatment)	1 and 3	-3.78	0.0001
C (treatment)	2 and 3	-3.66	0.0001

(inter-quartile range 5–12), while the other categories were rated lower; “Support from colleagues”, median 1 (inter-quartile range 0–3); “Philosophical opposition”, median 0 (inter-quartile range 0–3); “Client dissatisfaction”, median 2 (inter-quartile range 0–4); “Interference”, median 0 (inter-quartile range 0–2). Trainees reported perceptions of barriers similarly across professions. Report of barriers also did not differ across gender, age, or length of time worked in the health service. The above findings are consistent with the expected improvement in CBT competence following training and in terms of learner satisfaction.

Generalization Questionnaire

Contrary to expectations, however, surprisingly favourable findings were obtained in terms of the reported transfer of training. A McNemar test was run to evaluate differences between the paired scores on the first part of Section B (Generalization) of the Generalization Questionnaire. Table 3 presents the data on trainees’ use of CBT assessment instruments and interventions for the 3 months pre-training and for the 3 months following the course (Section B). These data indicate a statistically significant increase in the use of these strategies in the workplace post-training in all but one of the tools listed. “Graded exposure” was a strategy that was reported to be already widely used prior to training.

Data from 55 participants were available for the Generalization Questionnaire and pairwise comparisons were conducted using a Wilcoxon test. No significant differences were found on Sections A (“experience”: $z = -.38$ $p = .7$) and D (“support and participation” $z = -1.48$ $p = .14$). These results indicate that using CBT concepts (e.g. the use of the cognitive behavioural model, and support and participation in, for example clinical supervision and team meetings) did not change over the 3 months following the course. However, a significant difference was found on Section B (“generalization”: $z = -2.74$ $p < .006$) and Section C (“areas of functioning”: $z = -2.43$ $p < .015$), indicating an increase in both areas over time. We found that “generalization across persons” ($z = -2.23$ $p = .03$) and “response generalizations” ($z = -1.28$ $p = .2$) were not significantly altered from the post-training assessment, but that a significant improvement was found for “generalization across behaviours” ($z = -3.18$ $p < .001$). These findings indicate that the participants perceived themselves to have maintained their use of CBT methods during a 3-month follow-up period.

Table 3. Results of Generalization Questionnaire ($N = 55$)

Assessment instrument/CBT strategy	Probability (McNemar test)
Repeated measures (i.e. BDI, BHS)	.001
Cognitive behavioural formulation	.0001
Monitoring diaries	.001
Activity scheduling	.0001
Graded exposure	.5 ns
Thought catching	.0001
Socratic questioning	.001
Challenging negative automatic thoughts	.0001
Reality testing	.001
Downward arrow technique	.0001
Challenging dysfunctional assumptions	.0001
Relapse prevention strategies	.001

Discussion

In terms of the initial “reaction” evaluation, the course participants ($N = 55$) considered the course to be of high quality, relevant and felt that sufficient time was allocated for their active participation in experiential learning. The training was also considered to be acceptable and effective, receiving an overall endorsement of 75%.

The learning evaluation aspect also yielded promising results on the MCQ following training, and was maintained at a 3-month follow-up assessment. In addition, there were no significant differences found during the double baseline period ($N = 36$). Knowledge in CBT as measured by the VAT also improved significantly on identification of symptoms (Section A) and naming of appropriate CBT strategies (Section C), but no significant difference was found on Section B (problem identification), though there was a trend towards improvement in the relevant problem identification skills. The significant improvement in scores on sections A and C appears to reflect an increase in knowledge, one that was not obtained at the double baseline assessment. These findings indicate that several threats to the internal validity were not responsible for the improved knowledge scores. However, as cited earlier, measures of knowledge are not a reliable indicator of behaviour change. For instance, Reid and Whitman (1983) and Whitman, Scibak and Reid (1983) found that an increase in knowledge of treatment methods was not sufficient for trainees to apply what had been learned to the clinical setting. Nonetheless, it is clearly important for trainees to firstly acquire the necessary knowledge base before there is any application of the new clinical skills in the workplace.

The third aspect of the evaluation indicated that the largest barriers to implementing what had been learned to the workplace were the “institutional constraints”. These findings are similar to those of Backer, Liberman and Kuehnel (1986), Baldwin and Ford (1988) and Goldstein (1986), who found that work environment issues were the main factors affecting trainees’ ability to transfer training to the work setting. However, other dimensions of the BARCQ were perceived more favourably, which may explain the relatively good level of generalization obtained.

The training appeared to result in the generalization of knowledge and skills to the workplace, as measured by the Generalization Questionnaire at a 3-month follow-up. The results address the question posed by Bootzin and Ruggill (1988): “Do trainees retain their skills over time or

do they revert to idiosyncratic styles?" (p. 707). The results of this study indicate that trainees did retain their skills over time. Overall, therefore, we obtained evidence that the brief shared learning in CBT achieved the expected reaction, learning and generalization outcomes, the latter contrary to expectations.

Methodological critique

Although the results appear favourable, caution must be taken in their interpretation due to the incomplete data set, particularly from post-assessment to follow-up and the reduced data available for the VAT. In addition, as no direct measure of behaviour was used, the possibility of self-report biases cannot be discounted.

The lack of ability to identify the specific problem (section B) may have occurred for a number of reasons. It may simply be that the training did not enable trainees to achieve this skill. Alternatively, trainees may lack confidence in identifying disorders. As the participants were predominantly nurses and diagnosis is not part of their role, this may have added to a lack of confidence in labelling the problem as Panic Disorder, as opposed to stipulating it was an anxiety disorder. Another explanation may be that the measure lacks sensitivity: with a scoring of 0, 1, 2, the margin for improvement is small.

No direct observation of behaviour change following the training was used (e.g. submission of audio or videotapes of clinical practice). Similarly, no objective data were collected on any improvements to the work environment following training, and there was also no direct measure of benefits to patients. This latter area of concern could be rectified by asking patients to complete a measure of their satisfaction, and comparisons could be made using pre and post-treatment measures used in routine clinical practice, such as the Beck Depression Inventory (Beck, Steer, & Garvin, 1988) and the Beck Anxiety Inventory (Beck, Epstein, Brown, & Steer, 1988). These weaknesses could explain the unexpectedly favourable account of generalization as this was dependent upon the participants' self-ratings. As they were well aware of the aims of the training a bias may have been present, which direct observation or another objective assessment may have minimized.

The VAT is a more objective instrument, but it was altered half way through the study, due to the presence of a ceiling effect in Section C. Although the changes were necessary, the alterations reduced the amount of data available for analysis, possibly explaining why the results from Section B did not reach significance.

Conclusions

Research into the effectiveness of CBT training has been limited in number and methodological rigour. The present study appears to make a useful contribution to this sparse literature by providing a controlled, multiple measures/methods evaluation of shared learning for ($n = 90$) mental health professionals indicating that brief shared learning in CBT can be effective.

However, a major limitation of this study is the lack of direct measurement of the generalization of the training to the benefit of patients. Analysis of this form of generalization is an area frequently neglected by researchers. Ultimately, to be judged as successful, training in evidence-based practices must generalize to the benefit of patients. Researchers should therefore attempt to establish the extent to which training results in better clinical outcomes. For instance, Milne et al. (1999), in a study of CBT training, used the Coping Responses Inventory

(CRI: Moos, 1990) to measure individual patient's personal coping strategies. However, theirs was not a controlled evaluation.

Finally, the definition and measurement of independent variables in the training literature is rare. If researchers can ensure the validity of the independent variable (i.e. training) then replication is possible and we can be satisfied that the independent variable alone is responsible for the results, as opposed to other factors. Milne (2000) recommended this type of assessment, referred to as a "manipulation check". Waltz, Addis, Koerner and Jacobson (1993) argued that manipulation checks are crucial for high-quality research. They added that it is difficult to draw firm conclusions about the effects of the independent variable if a manipulation check is not carried out. They recommended that researchers assess competence through independent measurement, which should then be rated by an independent assessor, rather than making the assumption that level of training and experience determines competence.

Improved evaluations meeting these criticisms will help to demonstrate the value of CBT and provide a firm evidence base for a national programme of training. As illustrated in the present study, such training will help to address critical skill gaps.

Acknowledgements

We are grateful to the trainees for their co-operation with our assessment demands. Thanks to Vince Price, Esther Cohen-Tovee, Catriona Gray and Jennie Dickerson for their contributions to the teaching and to Julie Dodd for her help in preparing the manuscript.

References

- ALBERTS, G., & EDELSTEIN, B. (1990). Therapist training: A critical review of skill training studies. *Clinical Psychology Review*, *10*, 497–311.
- BACKER, T., LIBERMAN, R., & KUEHNEL, T. (1986). Dissemination and adoption of innovative psychosocial interventions. *Journal of Consulting and Clinical Psychology*, *54*, 111–118.
- BALDWIN, T. T., & FORD, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*, 63–105.
- BECK, A. T., EPSTEIN, N., BROWN, G., & STEER, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*, 893–897.
- BECK, A. T., STEER, R. A., & GARVIN, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*, 77–100.
- BOOTZIN, R. R., & RUGGILL, J. S. (1988). Training issues in behaviour therapy. *Journal of Consulting and Clinical Psychology*, *56*, 703–709.
- CORRIGAN, P. W., KARATANINI, W. Y., & PRAMANAN, W. (1992). Staff perceptions of barriers to behaviour therapy at a psychiatric hospital. *Behaviour Modification*, *16*, 132–144.
- DAVIS, J. R., RAWANA, E. P., & CAPPONI, D. R. (1989). Acceptability of behavioural staff management techniques. *Behaviour Residential Treatment*, *4*, 23–44.
- DEPARTMENT OF HEALTH (1999). *National Service Framework for Mental Health*. London: Department of Health.
- FREIHEIT, S. R., & OVERHOLSER, J. C. (1997). Training issues in cognitive-behavioural psychotherapy. *Journal of Behaviour Therapy and Experimental Psychiatry*, *28*, 79–86.
- GOLDSTEIN, I. L. (1986). *Training in organizations: Needs assessment, development and evaluation*. Monterey, CA: Brooks/Cole.
- JAHN, E. (1998). Current issues in staff training. *Research in Developmental Disabilities*, *19*, 73–87.

- KING, M., DAVIDSON, O., TAYLOR, F., HAINES, A., SHARPE, D., & TURNER, R. (2002). Effectiveness of teaching general practitioner skills in brief cognitive behavioural therapy to treat patients with depression: Randomized controlled trial. *British Medical Journal*, *324*, 947–950.
- MACA (1999). *First national General Practitioner survey of mental health in primary care*. London: Mental Aftercare Association.
- MILNE, D. L. (1984). Skill evaluations of nurse training in behaviour therapy. *Behavioural Psychotherapy*, *12*, 142–150.
- MILNE, D. L. (2000). Evaluation in mental health nursing. In R. Newell & K. Gournay (Eds.), *Mental health nursing: An evidence based approach* (pp.121–144). Edinburgh: Churchill Livingstone.
- MILNE, D. L., BAKER, C., BLACKBURN, I., JAMES, I., & REICHEL, K. (1999). Effectiveness of cognitive therapy training. *Journal of Behaviour Therapy and Experimental Psychiatry*, *30*, 81–92.
- MILNE, D. L., GORENSKI, O., WESTERMAN, C., LECK, C., & KEEGAN, D. (2000). What does it take to transfer training? *Psychiatric Rehabilitation Skills*, *4*, 259–281.
- MILNE, D. L., JAMES, I. A., KEEGAN, D., & CHALK, A. (2001). General utility interactional device for evaluating leadership (GUIDE). *Clinical Psychology and Psychotherapy* (in press).
- MILNE, D. L., KEEGAN, D., WESTERMAN, C., & DUDLEY, M. (2000). Systematic process and outcome evaluation of brief staff training in psychosocial interventions for severe mental illness. *Journal of Behaviour Therapy and Experimental Psychiatry*, *31*, 87–101.
- MILNE, D. L., & NOONE, S. (1996). *Teaching and training for non-teachers*. Leicester: BPS Books.
- MOOS, R. H. (1990). Coping Responses Inventory. In D. L. Milne, *Assessment: A mental health portfolio*. Windsor: NFER-Nelson.
- MYLES, P. J., & MILNE, D. L. (2001). *The development of two new instruments for the evaluation of foundation level training in cognitive behaviour therapy*. Strathclyde University; Glasgow: Annual Conference for the British Association of Behavioural and Cognitive Psychotherapy.
- OLIVER, J. R., & FLEMING, R. K. (1997). Applying within-subject methodology to transfer of training research. *International Journal of Training and Development*, *1*, 173–180.
- REID, D., & WHITMAN, T. (1983). Behavioural staff management in institutions: A critical review of effectiveness and acceptability. *Analysis and Intervention in Developmental Disabilities*, *3*, 131–151.
- ROTH, A., & FONAGY, P. (1996). *A critical review of psychotherapy research*. London: Guilford Press.
- SHAW, B. F., & WILSON-SMITH, D. (1988). Training therapists in cognitive-behaviour therapy. In C. Perris & I. M. Blackburn (Eds.), *Cognitive therapy in practice*. Heidelberg: Springer Verlag.
- STRUPP, H. H., BUTLER, S. F., & ROSSER, C. L. (1988). Training in psychodynamic therapy. *Journal of Consulting and Clinical Psychology*, *156*, 689–695.
- WALTZ, J., ADDIS, M. E., KOERNER, K., & JACOBSON, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*, *61*, 620–630.
- WHITMAN, T. L., SCIBAK, J. W., & REID, D. H. (1983). *Behaviour modification with the severely and profoundly retarded*. New York: Academic Press.
- WILLIAMS, R. M., MOOREY, S., & COBB, J. (1991). Training in cognitive-behaviour therapy: Pilot evaluation of a training course using the cognitive therapy scale. *Behavioural Psychotherapy*, *19*, 373–376.
- WRIGHT, J., HORLICK, S., BOUCHARD, C., MATHIEU, M., & ZEICHNER, A. (1977). The development of instruments to assess behaviour therapy training. *Journal of Behaviour Therapy and Experimental Psychiatry*, *8*, 281–286.
- YOUNG, J., & BECK, A. T. (1980, 1988). *Cognitive Therapy Scale: Rating manual*. Unpublished manuscript. Centre for Cognitive Therapy, University of Pennsylvania, Room 602, 133 South 36th Street, Philadelphia 19104, USA.