A qualitative comparison of cognitive-behavioural and evidence-based clinical supervision

Derek L. Milne[1]*, Robert P. Reiser[2], Tom Cliffe[1], Lauren Breese[1], Annabel Boon[1], Rosamund Raine[1] and Phillippa Scarratt[1]

[1]*Newcastle University, UK*
[2]*Palo Alto University, CA, USA*

Abstract. Despite the acknowledged importance of clinical supervision, controlled research is minimal and has rarely addressed the measurement or manipulation of clinical supervision, hampering our understanding and application of the different supervision methods. We therefore compared two related approaches to supervision, cognitive-behavioural (CBT) and evidence-based clinical supervision (EBCS), evaluating their relative effectiveness in facilitating the experiential learning of one supervisee. Drawing on a multiple-baseline $N = 1$ design, we gathered mostly qualitative data by means of an episode analysis, a content analysis, a satisfaction questionnaire, and interviews with the supervisor and supervisee. We found that the EBCS approach was associated with higher supervision fidelity and increased engagement in experiential learning by the supervisee. This case study in the evaluation of supervision illustrates the successful application of some rarely applied qualitative methods and some potential supervision enhancements, which could contribute to the development of CBT supervision.

Key words: CBT supervision, evidence-based clinical supervision, qualitative evaluation.

Introduction

Clinical supervision is a necessary part of modern healthcare (DoH, 2004), intrinsic to professional development and regulation [e.g. British Association for Behavioural and Cognitive Psychotherapy (BABCP); Latham, 2006]. The reasons for its importance are numerous, including enhancing clinical outcomes and guiding therapists (supervisees), all leading ultimately to the promotion of safe and effective practice (Falender & Shafranske, 2004). For the UK's National Health Service (NHS), supervision serves to promote a 'high-quality of practice', as it 'will encourage reflective practice' (DoH, 2004, p. 35).

However, the term 'supervision' covers a wide diversity of practices, from personal growth (e.g. Hawkins & Shohet, 2000) to systematic professional development (e.g. Schoenwald

---

et al. 2009). This reflects equally diverse ways of defining clinical supervision, from the inclusive (Bernard & Goodyear, 2004) to the exclusive (Milne, 2007). The general lack of precision in defining supervision, which unfortunately includes the NHS's own brief and limited definition (DoH, 1993), has the advantage of feasibility, but it hampers best practice, supervisor training, and research. For example, practitioners are unclear about what exactly supervision means, and what it is supposed to achieve (Lister & Crisp, 2005). Similarly, in a scrutiny of supervision within a sample of 32 clinical trials, Roth *et al.* (2010) noted that information about the supervision that took place within these trials was presented inconsistently and sketchily. Because of this general lack of precision, the present study adopts the empirically derived definition in Milne (2007), which is essentially: 'The formal provision, by approved supervisors, of a relationship-based education and training that is work-focused and which manages, supports, develops and evaluates the work of colleagues' (p. 3).

Rationale for the present study

These problems with definition are compounded by the weak status of research on supervision, which is scant and embryonic, while also featuring vague or absent conceptualization, limited rigour and poor measurement (Ellis *et al.* 1996), a situation that has not improved (Wheeler & Richards, 2007; Ellis *et al.* 2008). To rectify matters, Falender *et al.* (2004) urged that 'A range of research procedures should be employed, including, for example, self-report, experimental, single-subject repeated measures, qualitative' (p. 775). A related suggestion, made by successive researchers, is to improve the way that we attempt to measure supervision (e.g. Ellis & Ladany, 1997). For example, Watkins (1998) stated that 'In recent years, several supervision publications have emphasized one point vigorously: more valid, reliable, supervision-specific measures are needed to advance research efforts' (p. 94). Watkins (1997) and Falender *et al.* (2004) also viewed qualitative methods as a promising option. Watkins (1998) has clarified the way that qualitative research could contribute to our understanding. He suggests that we require: 'A greater focus on the behaviour of the supervisor, examining what actually happens in supervision, what supervisors and supervisees actually do ... with multiple indices to measure supervision process and outcome ... longitudinal studies ... observable behavioural data' (p. 96).

Exemplifying this kind of approach to qualitative research, Milne *et al.* (2008) used a single subject ($N = 1$) multiple baseline design to compare two micro-analytical approaches to evaluating supervision: a quantitative momentary time-sampling methodology and an episode-based qualitative analysis (based on Ladany *et al.* 2005). As predicted, significant differences were observed in specific categories of supervisory effectiveness in the intervention phase, as indicated by both types of analysis.

Another way to advance supervision research is to compare different approaches, but such comparative evaluations are rare, by contrast with evaluations of the effectiveness of single approaches, such as CBT supervision (Milne & James, 2000; Townend *et al.* 2002; Wheeler & Richards, 2007; Milne *et al.* 2010). In a rare example of a comparative evaluation, Bambling *et al.* (2006) conducted a random controlled trial comparing CBT and psychodynamic supervision, reporting no clinical difference (measured in terms of patients' scores on a depression measure at the end of an eight-session treatment period). However, analysis of the clinical outcomes of patients who received a problem-solving therapy without

supervision, suggested that supervision of either kind improved patient outcome. In a second comparative evaluation, Uys *et al.* (2005) found that the two supervision approaches that they compared both produced significantly improved supervisee ratings of supervision, but that neither approach was superior (i.e. a developmental model, and one based on Holloway's 1995 matrix model).

In summary, the rationale for the present study is to try and enhance supervision research by addressing some of the prevailing concerns, such as the need for supervision-specific measures, qualitative approaches, and comparisons between different approaches. This study has the potential to make an important contribution to the definition and practice of CBT supervision, benefitting researchers, supervisors and clinicians.

In response to this parlous situation within supervision research, the present study adopts an $N = 1$ multiple baseline design to analyse the qualitative outcomes of CBT supervision, in comparison to Evidence-Based Clinical Supervision (EBCS), a science-informed approach to conducting and evaluating supervision (Milne, 2009). The EBCS approach supplements CBT supervision by incorporating a developmentally and educationally informed framework (e.g. explicit educational needs assessment), attending systematically to the affective aspects of therapy and supervision (e.g. feeling reactions to the experiential learning exercises within supervision). EBCS is also based on an integrated programme of research, systematic supervisor training, and the operationalization of EBCS through a supervision manual. Specific procedural overlaps and distinctions between the two approaches are provided within Table 1.

Predictions

This paper builds on the reviewed studies above by reporting a comparative qualitative evaluation. We predicted that:

  (I) We would be able to reliably introduce EBCS and CBT supervision conditions (i.e. fidelity would be demonstrated).
 (II) EBCS supervision would demonstrate stronger effects, in terms of improvements in the supervisee's engagement in experiential learning.
(III) Supervisees would regard the CBT and EBCS approach as equally acceptable, based on their satisfaction data.

Method

Design

In order to compare the two approaches to supervision, we utilized a multiple baseline across participants, alternating treatment design. This $N = 1$ design is considered appropriate for this kind of comparison (Oliver & Fleming, 1997; Borckardt *et al.* 2008), especially given that both approaches were still in need of systematic operationalization, and that only a few rigorous $N = 1$ evaluations have been conducted (Milne, 2008).

In relation to predictions I (demonstrate fidelity) and II (demonstrate EBCS improves engagement in experiential learning), there were multiple (i.e. A-B-A and A-B-A-B) experimental phases across the three clients. For example, the A-B-A design alternated the

Table 1. *A comparison of key elements of CBT† and evidence-based clinical supervision* (*EBCS*)

	CBT	EBCS
Session structure		
Prioritization and discussion of agenda	*	*
Bridge from previous session: inquiry about previous supervision session and previously supervised cases	*	*
Assignment and review of homework	*	*
Supervisor provides capsule summaries	*	*
Supervisor elicits feedback from therapist	*	*
Supervisor gives feedback to therapist	*	*
Makes primary use of review of audio or videotapes in supervision, and, utilizes standardized instruments to assess competence	*	*
Session content		
Teaches cognitive therapy individual case conceptualization, structuring of therapy sessions, and specific cognitive and behavioural techniques	*	*
Session process		
Takes collaborative approach	*	*
Takes an empirical stance with therapist (finding out what works)	*	*
Utilizes guided discovery and Socratic questioning	*	*
Addresses the therapist's personal issues that interfere with therapy, therapist's maladaptive beliefs about therapy, interpersonal issues in supervision	*	*
Utilizes experiential exercises and learning experiments	*	**
Utilizes multimodal methods of teaching including reading, modeling, live demonstrations, role-play rehearsal and audio or videotapes	*	**
Focus on emotional experiencing and affective engagement in the session	–	**
Higher levels of challenging guided experiential learning with slightly higher levels of anxiety or discomfort	–	**
Utilizes an explicit experiential learning model based on Kolb (1984)	–	**

–, Absent; * common elements; ** enhanced in EBCS.
† Based on Padesky (1996), Liese & Beck (1997).

two approaches to supervision, starting with CBT supervision in the baseline (A) phase, switching to EBCS for the intervention phase (B), then reverting to the baseline condition once more. This provided a test of whether the two approaches could be introduced and withdrawn successfully. These phases included 37 consecutive, audio-taped sessions of supervision, taking place over an 11-month period. Three of the authors (L.B., A.B., P.S.) scrutinized all 37 sessions within the present study.

Instruments

The first measure evaluated predictions I and II, and was based on direct observation. This was the task analysis approach called 'episode analysis' (Ladany *et al.* 2005), which these authors argue enables the supervision process to be segmented into 'meaningful chunks' (p. 6), facilitating our understanding of this process and the supervisees' development. An episode consists of 'not only what is specifically discussed in supervision (the content), but also the
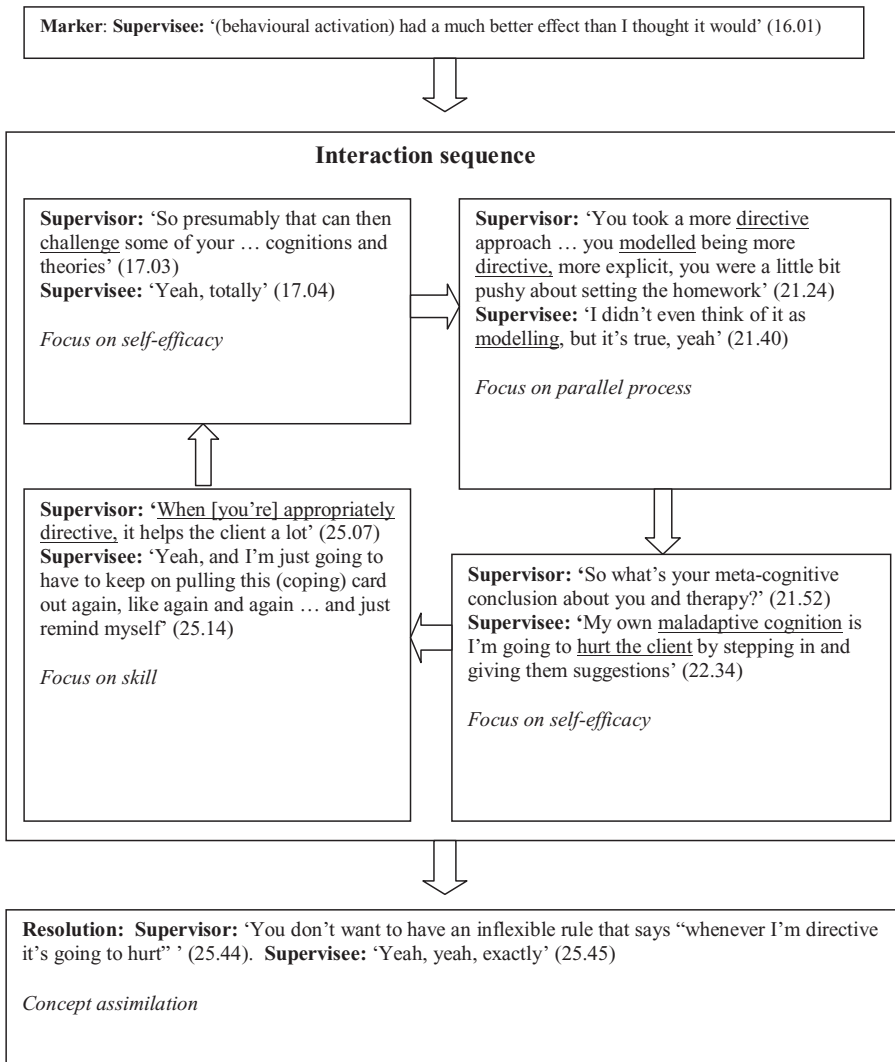
Marker: Supervisee: '(behavioural activation) had a much better effect than I thought it would' (16.01)

Interaction sequence

Supervisor: 'So presumably that can then <u>challenge</u> some of your … cognitions and theories' (17.03)
Supervisee: 'Yeah, totally' (17.04)

Focus on self-efficacy

Supervisor: 'You took a more <u>directive</u> approach … you <u>modelled</u> being more <u>directive</u>, more explicit, you were a little bit pushy about setting the homework' (21.24)
Supervisee: 'I didn't even think of it as <u>modelling</u>, but it's true, yeah' (21.40)

Focus on parallel process

Supervisor: '<u>When [you're] appropriately directive</u>, it helps the client a lot' (25.07)
Supervisee: 'Yeah, and I'm just going to have to keep on pulling this (coping) card out again, like again and again … and just remind myself' (25.14)

Focus on skill

Supervisor: 'So what's your meta-cognitive conclusion about you and therapy?' (21.52)
Supervisee: 'My own <u>maladaptive cognition</u> is I'm going to <u>hurt the client</u> by stepping in and giving them suggestions' (22.34)

Focus on self-efficacy

Resolution: Supervisor: 'You don't want to have an inflexible rule that says "whenever I'm directive it's going to hurt" ' (25.44). Supervisee: 'Yeah, yeah, exactly' (25.45)

Concept assimilation

Fig. 1. Illustrative episode from one supervision session.

types of sequential, interpersonal behaviours that can effect change' (Ladany *et al*. 2005, p. 6). Three elements compose an episode: a marker event (e.g. a question raised by the supervisee about the relevant knowledge base); an interaction sequence (various supervisor interventions, such as questioning and informing); and a resolution (e.g. a plan of action). Figure 1 illustrates this approach with an episode from a CBT supervision session (the figures in parentheses represent the time, in minutes and seconds; while the underlined words illustrate what was included in the content analysis).

The second measure was a content analysis, which was undertaken using transcripts of each observed episode. This allowed us to compare those episodes occurring within the two

Table 2. *Content analysis of all observed supervision episodes* (N = 31)

Type of supervision	Low fidelity content (frequency of utterances)	High fidelity content (frequency of utterances)
CBT	• Insight (3) • Counselling focus (32) (i.e. staying in present, Rogerian, early experiences, emotional experiences, exploring feelings, non-directive)	• Structuring session [e.g. agenda-setting (7)] • Specific cognitions identified (19) • Modelling (4) • Collaboration (4)
EBCS	• Counselling focus (5)	• Behavioural experiments (11) • Feeling reactions (73) • Role-play (20) • Video material (5) • Challenging (26)

EBCS, Evidence-based clinical supervision.

supervision approaches, CBT and EBCS, in a systematic and replicable manner (Bryman, 2004), providing a finer-grained qualitative method for testing the first two predictions. Coding categories were developed empirically by the first author, based on the manifest content of these episodes, following the procedure reported by Papworth *et al.* (2009) and from the defining features of CBT and EBCS, as described above (and see Table 1). This requires an iterative approach to be followed, in which the initial coding categories were revised as necessary with each successive episode, so that mutually exclusive categories of speech (i.e. utterances) were developed, as listed in Table 2. Once these categories were clarified they were then counted, to yield the frequencies of each type of utterance. Therefore, there was a minor quantitative aspect to the content analysis (this has been termed the 'hybrid' approach; Pistrang & Barker, 2010). The words spoken (i.e. utterances) by the supervisor and supervisee were treated together, and similarly no distinction was drawn between the detailed nature, quality, or other aspects of what was said: we simply listed and counted all utterances belonging to the identified categories.

The third and final prediction was assessed with a supervisee feedback and evaluation instrument, REACTS [Rating of Experiential learning And Components of Teaching & Supervision; Wilson, 2007 (note this instrument has not been published in a peer-reviewed journal)]. REACTS is an 11-item, one-page-long, paper-and-pencil rating of supervision by the supervisee, designed in part to assess Proctor's (1986) 'normative' and 'restorative' aspects of supervision (through items on the structure/resources: namely the frequency and duration of supervision sessions; and content that includes both management issues and the provision of emotional support). However, REACTS mainly focuses on the 'formative' aspect of supervision (i.e. educative function), by listing Kolb's (1984) learning modes (i.e. experiencing, reflecting, conceptualizing, experimenting, planning). An example item (number 5) is: 'I was able to recognize relevant feelings, becoming more self-aware (e.g. role-play helped me to express emotion)'. The five-point rating scale ranged from 'strongly agree' to strongly disagree' (with a 'not applicable' option), giving a score range of 8–45 (there are nine rated items), where higher scores represent greater supervisee/supervisor satisfaction and learning. REACTS also includes a 'Helpful aspects of supervision' item (following Llewellyn, 1988) to collect qualitative data, and a final item inviting any further comments. It can be

completed within 5 min. REACTS has good test–retest reliability ($r = 0.96$, $p = 0.0001$) and high internal consistency (Cronbach's $\alpha = 0.94$, $p = 0.001$).

The fourth and final measure was an interview, conducted independently with the supervisor and the second supervisee by a third-party interviewer, in order to assess the fidelity of the CBT and EBCS interventions. This was an *ad-hoc*, semi-structured interview, lasting up to 1 hour, intended to assess fidelity as defined by Bellg *et al.* (2004). It was developed following feedback from the supervisor, and consisted of a social desirability rating (made at the start and end of the interview by the interviewer) and six open-ended questions, each with prompts. For example, question 2 asked: 'What happened during evidence-based clinical supervision – what kind of supervision was it?' Prompts included asking what proportion of supervision sessions were made up of didactic and experiential work; and whether the EBCS guidelines and other supporting materials were ever used. There were no formal psychometric data for the interview. (Copies of all four instruments are available from the first author.)

Participants

The participants were one male consultant (i.e. the supervisor of the supervisor), one male supervisor, two female therapists (i.e. supervisees) and three clients (two males, one female). The consultant was a doctoral-level, fully accredited clinical psychologist, with over 30 years of experience in providing supervision and training to psychologists and other mental health professionals. The supervisor was also an accredited clinical psychologist, with 20 years of clinical experience, and was a training clinic director with over 6 years of experience in providing doctoral-level training. The first supervisee participated for prediction I. She was a second-year Ph.D. student. For predictions II and III there was a second supervisee, who was a post-doctoral student on a 1-year internship. Both supervisees were supervised by the same supervisor, successively. The adult, outpatient clients were two males and a female, who presented with anxiety or depression.

This was a convenience sample, having been selected on the pragmatic grounds that the supervisor had invited the consultant to engage in a collaborative research study during a representative period of clinic activity. Similarly, the supervisees were those who were currently working with the supervisor at the time of the study; and the three clients were those being seen by the second supervisee (i.e. the therapist) at the time of the study, and who completed all study phases (i.e. a further three clients discontinued therapy or did not consent to participate). In essence, therefore, this was a routine or naturalistic service evaluation, triggered by the consultant's involvement in the present research project. There were, therefore, no exclusion criteria and in general the sample was considered to be representative of the supervisees and clients within this clinic. The study was approved by the relevant University Institutional Review Board in the USA; and by the Research and Development Department of the consultant's employing NHS Trust in the UK. Informed consent was obtained from the supervisees and the clients.

Procedure

The study site was a community-based psychology training clinic in the USA, a clinic for adults presenting with complex mental health problems. Supervisor training took place

over five consecutive weeks immediately prior to the initial study baseline, following the same consultancy procedure that will be described below. During the training, baseline and experimental phases, consultancy entailed fortnightly, phone-based, hour-long reviews of the preceding week's tape recorded supervision (i.e. 'supervision-of-supervision'). Consultancy involved didactic work, including reference to a supervision manual, to supervision guidelines and to other supportive materials (such as a client experiencing scale, a supervisee's learning outcomes list, and illustrative video material; Milne, 2009). This didactic aspect was supplemented by experiential work, primarily revolving around corrective feedback to the supervisor. This was based on the consultant's completion of a supervisory competence rating scale (SAGE; Milne *et al*. 2011) and the supervisee's completion of her satisfaction form (REACTS); and experiential work such as educational role-plays and behavioural rehearsal. Consultancy within both approaches was intended to be equivalent, but with the appropriate differences of emphasis (i.e. to encourage adherence to the respective CBT and EBCS supervision approaches). All supervision sessions were recorded on audiotape by the supervisor, to permit the qualitative analyses. A total of 37 supervision sessions were recorded, each lasting approximately 1 hour and taking place in the supervisor's office. This sample was selected as it included all recorded supervision sessions within the study period in which there was a discussion of one or more of the three clients, during the entire 11-month study period.

Regarding the observation and recording of the episodes, for training purposes, one tape was initially analysed simultaneously by three coders. Each episode that we observed was transcribed and the key utterances of the supervisor and supervisee were entered into the Ladany *et al*. (2005) episode framework, to show the elements of an episode. Episodes were coded following a manual, developed specially for the present study and intended to operationalize the Ladany *et al*. (2005) approach. This particular supervision session was chosen as it had previously been coded by the first author (a supervision researcher, experienced in using direct observation), and it represented a clear example of a critical event, as defined by Ladany *et al.* (2005). It had been identified by the first author in routine consultancy, mid-way through the study. During training the tape was paused after each of the identified elements of the episode, and discussions between the three coders occurred until agreement was reached. This process was continued until the end of the tape (64 min), following the coding manual instructions. A second tape was then coded independently, to assess the observer's inter-rater agreement. Specifically, any markers and resolutions were identified, with both the time and corresponding verbal utterances being recorded. These data enabled agreement to be assessed using the exact % agreement method:

$$\text{Total \% agreement} = \frac{\text{number of agreements} - \text{number of disagreement}}{\text{total number of ratings}} \times 100.$$

Both tapes were representative of our dataset, containing episodes. After training, the three coders worked independently. When listening to an audiotape, any identified markers were recorded, alongside the corresponding time of the utterance. If the marker developed into a full episode it was documented; when a marker did not resolve it was not included in the study.

In terms of the remaining two measures, for prediction III only the second supervisee was requested to complete the satisfaction questionnaire after each supervision session (i.e.

Table 3. *Comparison of the frequency and nature of episodes that were observed in CBT and evidence-based clinical supervision* (*EBCS*)

Supervision approach	Frequency of episodes (% of supervision sessions with episodes)			Nature of episodes (all clients)		
	Clients discussed in supervision				Interaction	
	1	2	3	Markers	sequences	Resolutions
CBT	4 (50%) Total 16 (64%)	7 (78%)	5 (63%)	Need for guidance or corrective feedback (14), direction with client (2)	Focus on: skill (9), therapeutic process (4), alliance (3)	Skill enhancement (8), knowledge (4), self-awareness (3),
EBCS	7 (54%) Total 15 (66%)	5 (100%)	3 (43%)	Need for guidance or corrective feedback (14), process concerns (1), direction for client (1)	Focus on: skill (8), therapeutic process (6), exploration of feelings (1)	Skill enhancement (8), self-awareness (5), alliance (2), knowledge (1)

%, Proportion of episodes per taped supervision session.

pragmatically, this was all that was necessary to test this prediction); and the interviews were conducted over the telephone, 1-month after the end of the $N = 1$ study.

Results

Prediction (*I*): We would be able to reliably introduce EBCS and CBT supervision conditions.

To check the fidelity of the two supervision approaches, we compared the relevant supervisor utterances within all 31 supervision episodes (see below for a summary of these episodes). In this content analysis we noted all specific instances of key terms within these episodes, as illustrated in Figure 1 (i.e. the underlined words, e.g. 'challenge', 'directive', 'modelled'). We then combined them into suitable categories, as set out in Table 3. Table 2 summarizes the most frequently occurring categories, and the ones that best differentiated the supervisor's utterances in the CBT and EBCS phases.

Table 2 also clarifies the extent to which the utterances were either inconsistent with the given supervision phase (i.e. 'low fidelity content', e.g. discussing 'insight' within CBT supervision); or were consistent with that approach (i.e. 'high fidelity content', e.g. discussing 'feeling reactions' within EBCS). This content analysis suggested that CBT was more cognitive, structured and collaborative. By contrast, EBCS episodes contained more experiential (behavioural and affective) material, were more challenging, and contained fewer counselling-focused utterances.

We next supplemented this content analysis with interview data. The interviews, conducted independently with the supervisor and the second supervisee, were both judged by the interviewer (from the interview) and the first author (from the recorded replies) to be low in social desirability. Both interviewees stated that EBCS was the more structured approach, being more agenda-driven. Moreover, they both identified EBCS as more affectively charged ('challenging', 'arousing', 'emotional') and more experiential ('role-plays'). This experiential emphasis was associated with the use of several materials from the EBCS manual (i.e. the client experiencing scale, learning outcomes list, supervision guidelines, emotions library, video material). When asked about the negative effects of EBCS, the supervisee noted that supervision sessions could be: 'rushed, as there was so much to fit in'. EBCS was also regarded as more taxing than CBT supervision, including 'raising anxiety levels'. The supervisor agreed that EBCS was 'challenging' and that he too experienced some anxiety. However, both construed this arousal as desirable and productive. The positive effects of EBCS were thought to revolve around the unprecedented attention that was accorded to the facilitation of experiential learning ('never really done experimenting before', 'directly challenged beliefs', 'deepening beliefs', 'enhanced learning'). Neither of the participants thought that the results obtained were due to factors other than EBCS, although each noted some other influences (e.g. the supportive relationship, the broad range of clients). In summary, the content analysis and interviews indicated that the two supervision approaches were implemented with fidelity, although the episode analysis suggested that the CBT supervision had high frequencies of counselling-style utterances, which might lead to questioning its fidelity.

Prediction (*II*)*:* EBCS supervision would demonstrate stronger effects, in terms of improvements in the supervisee's engagement in experiential learning.

As regards the relative effectiveness of EBCS over CBT supervision, assessed in terms of the supervisee's engagement in experiential learning, we found that 31 of the 37 tapes contained episodes, but (contrary to our prediction) that these were divided roughly equally between the CBT and EBCS phases of supervision (16 and 15, respectively). Table 3 shows that these episodes were also distributed across all three clients roughly equally (although discussion of client 2 was associated with the most episodes per session, in both phases). That is, episodes were observed on 64% (CBT supervision) and 60% (EBCS) of occasions, overall, but client 2 had the highest frequency of episodes in both phases (i.e. 78% and 100%).

It can also be seen from Table 3 that the nature of the episodes is comparable. Using the categories provided by Ladany *et al*. (2005), we found that the markers in both approaches were mainly concerned with the supervisee's need for guidance, leading in the main to the supervisor focusing on key skills, and resolving the episodes with the identification of skill developments. In summary, these episode data do not suggest that there is any difference in the effectiveness of the CBT and EBCS approaches.

The final prediction was that both forms of supervision would achieve similarly favourable qualitative feedback and evaluation ratings from the supervisee. A total of 20 REACTS forms were returned by the second supervisee, a 54% response rate. This predicted reaction was obtained, as the supervisee rated the supervisor very highly throughout the study: on all but seven items out of a possible 180 (i.e. 4%), the supervisor was credited with the maximum rating of 5 ('strongly agree'). Four of these lower ratings (i.e. a rating of 3 or 4) were made during CBT supervision, but there was no pattern in terms of the items receiving these

lower ratings. In summary, expressed as a percentage of the maximum possible supervisee satisfaction, both approaches received very high endorsement (98% for CBT, 99% for EBCS). This is equivalent to an overall rating of 'strongly agree' in both cases.

The qualitative item concerned with 'helpful events' elicited comments on all 20 completed REACTS forms. During CBT supervision these comments highlighted the value of planning what to do next in therapy (e.g. 'looking at the case from different angles'). In EBCS the helpful events that were noted were far more varied and numerous, embracing a wide range of experiential methods (e.g. role-plays, viewing video recordings of the therapy, and exploring feeling reactions). These methods were frequently associated with comments about effective treatment planning, enhanced self-awareness, and improved self-confidence. Last, the final REACTS item, 'Any other comments?' received only two responses: in the CBT instance the supervisee noted the need to allocate the time better, so that all clients could be discussed; in the EBCS instance, she noted that it had been 'difficult being admonished for my lack of professionalism'.

Discussion

Our first prediction was that the two supervision approaches, CBT and EBCS, could be reliably introduced. Such fidelity is a basic logical requirement for process-outcome research, as well as for increasing power (Lipsey, 1990). However, it is rare for studies in the supervision field to include demonstrations of fidelity (Watkins, 1997). Using a combination of content analysis and interviewing, we were able to provide evidence that the CBT and EBCS approaches could be introduced with reasonably good fidelity between the phases (i.e. A-B-A-B). To illustrate, this qualitative analysis suggested that CBT was more cognitive, structured and collaborative. By contrast, EBCS episodes contained more experiential (behavioural and affective) material, were more challenging, and contained fewer counselling-focused utterances. One significant discrepancy was the high frequency of 'counselling focus' in CBT supervision. It may be that this actually subsumes a high proportion of utterances that would be deemed perfectly appropriate within CBT supervision, because they reflected aspects of the supervisory alliance and a collaborative style (e.g. providing core conditions, such as empathic statements). Therefore, we suggest that this requires further investigation, ideally guided by a CBT supervision manual or similar operational statement. On the other hand, if we assume that the 'counselling focus' was indeed due to low fidelity to the CBT supervision approach, then the differential fidelity confounds our comparison: what we may have is a comparison between a high *vs.* a low fidelity intervention. It would not be surprising if this produced findings favouring the high fidelity approach. Again, we believe that the way forward is to operationalize CBT supervision, so that fidelity can be quantified and manipulated more precisely. As to why we found this differential fidelity, the supervisor was far more experienced in CBT supervision, so we believe that the explanation is that the supervisor was provided with individual training ('supervision-of-supervision') in EBCS using a systematically prepared training package including a clear operationalization (Milne, 2010), whereas there was no such package for CBT. There may also have been 'demand characteristic' and 'allegiance' effects in operation, as the supervisor and the first author were active research collaborators, developing the EBCS approach.

Our second main finding was that the CBT and EBCS approaches appeared to yield predictably different effects, as measured in terms of the supervisee's engagement in

experiential learning. The qualitative evaluation was based initially on the episode approach (Ladany *et al.* 2005), which indicated that both approaches were equally effective. A total of 37 tapes contained episodes, divided roughly equally between the CBT and EBCS phases of supervision (16 and 15, respectively). However, when these episodes were probed for their content, EBCS was shown to promote much more frequent engagement in experiential learning, in terms of the discussion of topics like behavioural experiments (11 instances), feeling reactions (73) and role-play (20) (see Table 2, high fidelity column).The final prediction was that both forms of supervision would gain similarly favourable satisfaction ratings from the supervisee, for which we found clear support, although this was possibly confounded because the supervisee rated the supervisor very highly throughout the study (98% for CBT, 99% for EBCS). We strongly suspect that a positive bias inflated these ratings, a recurring problem with satisfaction ratings that are provided to a supervisor. The qualitative item concerned with 'helpful events' elicited comments on all the REACTS forms. During CBT supervision these comments highlighted the value of planning what to do next in therapy (e.g. 'looking at the case from different angles'). As just noted in relation to the content analysis (Table 2), in EBCS the helpful events that were noted were far more varied and numerous, embracing a wide range of experiential methods (e.g. role-plays, viewing video recordings of the therapy, and exploring feeling reactions). These methods were frequently associated with comments about effective treatment planning, enhanced self-awareness, and improved self-confidence. Thus, these qualitative data seem to provide more useful discriminations between the two approaches than do the ratings, due to a ceiling effect (and to the effects of a 'grateful testimonial').

Critical review

We recognize a number of additional methodological weaknesses within the present analysis. These include the need to add complementary instruments, such as objective (quantitative) assessments of learning and of the transfer of the supervisees' learning to therapy and to refine training in EBCS. The concept of 'episode' may also merit refinement, given striking parallels with concepts such as 'sudden gains' (Tang & DeRubeis, 1999), 'milestones' (Shiffman *et al*. 2006), 'good moments' (Mahrer & Nadler, 1986) and 'achievements' (Barkham *et al.* 2010). Another weakness was our use of an experienced supervisor, in the sense that he would perhaps have an established allegiance to his usual approach (CBT supervision), or be less markedly influenced by training and consultancy, a common finding within therapist-training studies (Beidas & Kendall, 2010). A more substantive methodological issue concerns the differentiation of CBT supervision and EBCS, in that it could be argued that EBCS is simply CBT supervision done correctly (i.e. the argument that EBCS is not conceptually different from CBT supervision, even if there are differences in implementation). After all, the main accounts of CBT supervision advocate much that is in EBCS (e.g. Padesky, 1996; Liese & Beck, 1997), such as educational role-play. However, we believe that there are practical and theoretical differences between the two approaches: in addition to a different emphasis on some shared variables (e.g. EBCS stresses the behavioural and affective aspects of supervision), there do appear to be conceptually distinct aspects of EBCS, due to drawing on ideas about human development and learning from beyond the CBT supervision literature (Milne, 2009). Further, CBT supervision as implemented does not appear to match these theoretical accounts, either in the present study or in the other process evaluations that

exist (Milne, 2008), complicating differentiation. It appears that CBT supervision requires the kind of operationalization that has occurred with EBCS before a true comparison can be undertaken, and before appropriate conclusions can be drawn about distinctiveness, relative effectiveness, etc. Following a related example, where motivational interviewing was differentiated from acceptance and commitment therapy in terms of theoretical and practical comparisons (Bricker & Tollison, 2011), it would seem reasonable to refer to CBT and EBCS as 'complementary' approaches. Future research might draw on these points, to enable a more systematic comparison between CBT, EBCS and other supervision models. This may help to identify the key conceptual features, relevant instruments, and the most effective elements within such approaches (Kazdin, 1998).

Conclusions

This illustrative case study indicates the potential for qualitative methods to advance our understanding of CBT supervision. The present analysis was novel in clarifying qualitatively the fidelity and comparative effectiveness of two supervision approaches. However, improved $N = 1$ studies are required, before proceeding to large-group evaluations.

Declaration of Interest

None.

Recommended follow-up reading

Miller WR, Rollnick S (2009). Ten things that motivational interviewing is not. *Behavioural and Cognitive Psychotherapy*, **37**, 129–140.

References

Bambling M, King R, Raue P, Schweitzer R, Lambert W (2006). Clinical supervision: its influence on client-rated working alliance and client symptom reduction in the brief treatment of major depression. *Psychotherapy Research* **16**, 317–331.

Barkham M, Stiles WB, Lambert MJ, Mellor-Clark J (2010). Building a rigorous and relevant knowledge-base for the psychological therapies. In: *Developing and Delivering Practice-Based Evidence* (ed. M. Barkham, G. E. Hardy and J. Mellor-Clark), pp. 21–61. Chichester: Wiley.

Beidas RS, Kendall PC (2010). Training therapists in evidence-based practice: a critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice* **17**, 1–30.

Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, Ogedegbe G, Orwig D, Ernst D, Czajkowski S (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychology* **23**, 443–451.

Bernard JM, Goodyear RK (2004). *Fundamentals of Clinical Supervision*, 3rd edn. London: Pearson.

Borckardt JJ, Nash MR, Murphy MD, Moore M, Shaw D, O'Neil P (2008). Clinical practice as natural laboratory for psychotherapy research. *American Psychologist* **63**, 77–95.

Bricker J, Tollison S (2011). Comparison of motivational interviewing with acceptance and commitment therapy: a conceptual and clinical review. *Behavioural and Cognitive Psychotherapy* **39**, 541–559.

Bryman A (2004). *Social Research Methods.* Oxford University Press.

Cohen J (1988). *Statistical Power Analysis for the Behavioural Sciences*. New York: Lawrence Erlbaum.

DoH (1993). *A Vision for the Future*. London: Department of Health.

DoH (2004). *Organising and Delivering Psychological Therapies*. London: Department of Health.

Ellis MV, D'Iuso N, Ladany N (2008). State of the art in the assessment, measurement, and evaluation of clinical supervision. In: *Psychotherapy Supervision: Theory, Research and Practice* (ed. A. K. Hess, K. D. Hess and T. H. Hess), pp. 473–499. Chichester: Wiley.

Ellis MV, Ladany N (1997). Inferences concerning supervisees and clients in clinical supervision: an integrative review. In: *Handbook of Psychotherapy Supervision* (ed. C. E. Watkins), pp. 447–507. New York: Wiley.

Ellis MV, Ladany N, Krengel M, Schult D (1996). Clinical supervision research from 1981 – 1993: a methodological critic. *Journal of Counselling Psychology* **43**, 35–40.

Falender C, Cornish JAE, Goodyear R, Hatcher R, Kaslow NJ, Leventhal G,Shafranske E, Sigmon ST, Stoltenberg C, Grus C (2004). Defining competencies in psychology supervision: a consensus statement. *Journal of Clinical Psychology* **60**, 771–785.

Falender CA, Shafranske EP (2004). *Clinical Supervision: A Competency-Based Approach*. Washington, DC: American Psychological Association.

Hawkins P, Shohet R (2000). *Supervision in the Helping Professions: An Individual, Group and Organizational Approach*. Milton Keynes: Open University Press.

Holloway EL (1995). *Clinical Supervision: A Systems Approach*. Thousand Oaks, CA: Sage.

Kazdin AE (1998). *Research Design in Clinical Psychology*. Boston: Allyn Bacon.

Kolb DA (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice-Hall.

Ladany N, Friedlander ML, Nelson ML (2005). *Critical Events in Psychotherapy Supervision: An Interpersonal Approach*. Washington, DC: American Psychological Association.

Latham M (2006). Supervisor and training accreditation (Training Newsletter, February, p. 3). Accrington: British Association of Behavioural and Cognitive Psychotherapies (BABCP).

Llewellyn SP (1988). Psychological therapy as viewed by clients and therapists. *British Journal of Clinical Psychology* **27**, 105–114.

Liese BS, Beck JS (1997). Cognitive therapy supervision. In: *Handbook of Psychotherapy Supervision* (ed. C. E. Watkins), pp. 114–133. Chichester: Wiley.

Lipsey MW (1990). *Design Sensitivity*. London: Sage.

Lister PG, Crisp BR (2005). Clinical supervision in child protection for community nurses. *Child Abuse Review* **14**, 57–72.

Mahrer AR, Nadler WP (1986). Good moments in psychotherapy: a preliminary review, a list, and some promising research avenues. *Journal of Consulting and Clinical Psychology* **54**, 10–15.

Milne DL (2007). An empirical definition of clinical supervision. *British Journal of Clinical Psychology* **46**, 437–447.

Milne DL (2008). CBT supervision: from reflexivity to specialization. *Behavioural and Cognitive Psychotherapy* **36**, 779–786.

Milne DL (2009). *Evidence-Based Clinical Supervision: Principles and Practice*. Chichester: BPS/Blackwell.

Milne DL (2010). Can we enhance the training of clinical supervisors? A national pilot study of an evidence-based approach. *Clinical Psychology and Psychotherapy* **17**, 321–328.

Milne DL, James IA (2000). A systematic review of effective cognitive-behavioural supervision. *British Journal of Clinical Psychology* **39**, 111–127.

Milne DL, Lombardo C, Kennedy E, Freeston M, Day A (2008). Zooming in on clinical supervision. *Behavioural and Cognitive Psychotherapy* **36**, 619–624.

Milne DL, Reiser R, Aylott H, Dunkerley C, Fitzpatrick H, Wharton S (2010). The systematic review as an empirical approach to improving CBT supervision. *International Journal of Cognitive Therapy* **3**, 278–294.

Milne DL, Reiser R, Cliffe T, Raine R (2011). SAGE: preliminary evaluation of an instrument for observing competence in CBT supervision. *The Cognitive Behaviour Therapist*. Published online: 24 November 2011. doi:10.1017/S1754470X11000079.

Oliver JR, Fleming RK (1997). Applying within-subject methodology to the transfer of training research. *International Journal of Training and Development* **1**, 173–180.

Padesky CA (1996). Developing cognitive therapist competency: teaching and supervision models. In: *Frontiers of Cognitive Therapy* (ed. P. M. Salkovskis), pp. 266–292. London: Guilford Press.

Papworth MA, Milne DL, Boak G (2009). An exploratory content analysis of situational leadership. *Journal of Management Development* **28**, 593–606.

Pistrang N, Barker C (2010) Scientific, practical and personal decisions in selecting qualitative methods. In: *Developing and Delivering Practice-Based Evidence* (ed. M. Barkham, G. E. Hardy and J. Mellor-Clark), pp. 65–90. Chichester: Wiley.

Proctor B (1986). Supervision: a co-operative exercise in accountability. In: *Enabling and Ensuring* (ed. M. Marken and M Payne), pp. 21–34. Leicester: Leicester National Youth Bureau and Council for Education and Training in Youth and Community Work.

Roth AD, Pilling S, Turner J (2010). Therapist training and supervision in clinical trials: Implications for clinical practice. *Behavioural and Cognitive Psychotherapy* **38**, 291–302.

Schoenwald SK, Sheidow AJ, Chapman JE (2009). Clinical supervision in treatment transport: effects on adherence and outcomes. *Journal of Consulting and Clinical Psychology* **77**, 410–421.

Shiffman S, Scharf DM, Shadel WG, Gwaltney CJ, Dang Q, Paton Clark (2006). Analyzing milestones in smoking cessation: Illustration in a nicotine patch trial in adult smokers. *Journal of Consulting and Clinical Psychology* **74**, 276–285.

Tang T, DeRubeis RJ (1999). Sudden gains and critical sessions in cognitive-behavioural therapy for depression. *Journal of Consulting and Clinical Psychology* **67**, 894–904.

Townend M, Iannetta L, Freeston MH (2002). Clinical supervision in practice: a survey of UK cognitive-behavioural psychotherapists accredited by the BABCP. *Behavioural and Cognitive Psychotherapy* **30**, 485–500.

Uys LR, Minnaar A, Simpson B, Reid S (2005). The effect of two models of supervision on selected outcomes. *Journal of Nursing Scholarship* **37**, 282–288.

Watkins CE (1997). *Handbook of Psychotherapy Supervision*. Chichester: Wiley.

Watkins CE (1998). Psychotherapy supervision in the 21st century. *Journal of Psychotherapy Practice & Research* **7**, 93–101.

Wheeler S, Richards K (2007). The impact of clinical supervision on counsellors and therapists, their practice and their clients: a systematic review of the literature. *Counselling and Psychotherapy Research* **7**, 54–65.

Wilson M (2007). Can experiences of supervision be quantified? REACTS: a new tool for measuring supervisee's perceived satisfaction with clinical supervision [unpublished thesis] (available from the Psychology Department, Newcastle University).

Learning objectives

By studying this paper carefully, the reader will be able to:

(1) Summarize the argument for comparative evaluations.
(2) Discuss the criteria for distinguishing between two or more interventions.
(3) Evaluate the extent to which the present study represents a comparative evaluation of two forms of supervision.